国外农业科学经典著作译丛

数量遗传学导论

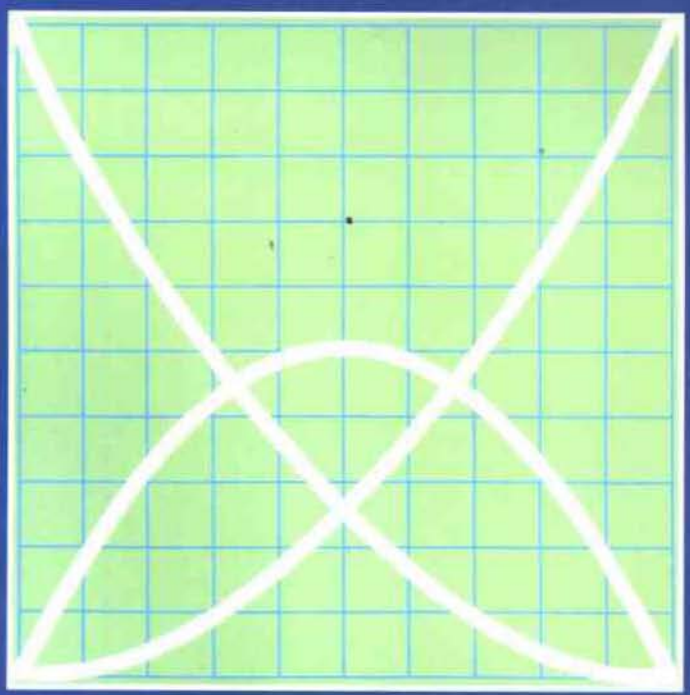
（第四版）

INTRODUCTION TO QUANTITATIVE GENETICS

FOURTH EDITION

[英] D.S.Falconer [美] T.F.C.Mackay 著

储明星 译 师守堃 校



中国农业科技出版社

责任编辑 冯凌云
封面设计 孙宝林

ISBN 7-80119-787-9/S·411
定价：80.00元

国外农业科学经典著作译丛

数量遗传学导论

（第四版）

[英]D.S. Falconer [美]T.F.C. Mackay 著

储明星 译 师守堃 校

中国农业科技出版社

图书在版编目（CIP）数据

数量遗传学导论（第四版）= Introduction to Quantitative Genetics（Fourth Edition）/（英）法尔康纳(Falconer,D.S.)，(美)麦凯（Mackay，T.F.C.）著；储明星译 .—北京：中国农业科技出版社，2000.5

（国外农业科学经典著作译丛）

ISBN 7-80119-787-9

Ⅰ.数… Ⅱ.储… Ⅲ.数量遗传学 Ⅳ.Q348

中国版本图书馆 CIP 数据核字（1999）第 12785 号

丛书策划	冯志杰
责任编辑	冯凌云
出版发行	中国农业科技出版社
	地址：北京海淀区白石桥路 30 号　邮编：100081
	电话：(010) 68919711；68975144；传真：68919698
经　销	新华书店北京发行所
印　刷	北京印刷学院印刷厂
开　本	787mm × 1092mm　1/16　印张：25.25
印　数	1 ~ 1 000 册　　字数：600 千字
版　次	2000 年 5 月第 1 版，2000 年 5 月第 1 次印刷
定　价	80.00 元

Introduction to Quantitative Genetics

(Fourth Edition)

D.S. Falconer
Formerly with Institute of Cell, Animal and Population Biology
University of Edinburgh

and

Trudy F.C. Mackay
Department of Genetics
North Carolina State University

Longman Group Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
United Kingdom
©D.S. Falconer 1960, 1981, 1989
This edition © Longman Group Limited 1996

This translation of Introduction to Quantitative Genetics, Fourth Edition, is published by arrangement with Addison Wesley Longman Limited.

内容简介

本书全面系统地介绍了数量遗传学的基本原理。全书由两大部分组成。第一部分包括1～5章，主要描述生物群体的遗传特性，被称为“群体遗传学”；第二部分包括6～21章，主要描述生物数量性状变异并引入与度量的遗传特性有关的概念，被称为“数量遗传学”或“统计遗传学”。

本书内容丰富，资料翔实，既可作为高等农业院校动植物遗传育种专业本专科生、研究生的参考教材，也可作为从事有关动植物科研、教学和生产人员的参考书。

中文版序言（一）

20世纪中叶，数量遗传学理论体系不断得到完善和发展，并指导动植物改良的实践取得了明显效果。Falconer教授所著的《数量遗传学导论》是该领域较早系统阐述数量遗传学的学术著作之一，自1960年问世以来已经多次修订，为数量遗传学的完善和发展起到了重要作用。

1996年出版的该书第四版是最新版本，内容较以往更为丰富，不仅涵盖了群体的基因频率变化、连续变异、变异分析、遗传参数估计、近交和杂交、自然与人工选择等经典数量遗传学的基本理论与方法，而且还吸收了数量遗传学领域的最新研究成果。该书第四版在对原有章节进行修订的同时，还增加了数量性状位点、数量遗传学的新理论和新方法、进化数量遗传学以及基因分子标记等较新颖的内容，使得数量遗传学以新的面貌展示在读者面前。

由中国农业科学院畜牧研究所储明星博士翻译的该书第四版中文本即将由中国农业科技出版社出版面世，这无疑是一件有意义的事情，相信该书在我国的翻译出版会对我国相关领域的科研、教学和生产起到促进作用。因此，我很高兴能为该书中文本作序。

吴常信

2000年1月于北京

中文版序言（二）

数量遗传学自20世纪中叶诞生以来获得了很大的发展，其理论体系不断得到完善和发展，并指导着动植物改良的实践，为人类的食物生产发挥了重要作用。这在D. S. Falconer所著《数量遗传学导论》中有着比较全面的反映。该书系统总结了半个多世纪以来数量遗传学的发展，展示了数量遗传学理论在动植物改良中的应用成果。自1960年问世以来，该书迄今已再版4次，重印10余次，不仅对数量遗传学理论有所发展，而且使其得到广泛传播，在数量遗传学乃至整个遗传学领域有着广泛的影响。由储明星博士翻译本书第四版中文本即将由中国农业科技出版社出版面世，这对从事遗传学，特别是数量遗传学乃至生物学的科研、教学人员来说无疑是一件有意义的事情。

该书作为数量遗传学领域的一部经典学术著作，理论系统性强是其一个突出特点。数量遗传学的理论基础最早是由Fisher等在20世纪20年代创立的，以后不断得到完善和发展。而Falconer所著的《数量遗传学导论》是该领域较早系统阐述数量遗传学的著作之一，自1960年问世以来已经多次修订，为数量遗传学的完善和发展起到了重要作用。该书的最新版更加系统和全面，不仅涵盖了群体的基因频率变化、连续变异、变异分析、遗传参数估计、近交和杂交、自然与人工选择等经典数量遗传学的基本理论与方法，而且还吸收了数量遗传学领域的最新研究成果，使本书更加系统和完善。

注重吸收新的研究成果，不断丰富和发展数量遗传学理论和方法是本书的第二个特点。在20世纪，科学技术的发展日新月异，数量遗传学也不例外。应当说，目前的数量遗传学理论体系自诞生以来，日趋完善，但另一方面，科学的发展又是无止境的。这一点通过《数量遗传学导论》一书的不同版本也可略见一斑。该书作者Falconer教授紧紧抓住数量遗传学发展的新动向，通过吸收新的研究成果完善并发展数量遗传学理论，也为数量遗传学这一学科保持旺盛的生命力起到了应有的作用。这也是本书具有极强生命力的原因所在。本书第四版在对原有的章节进行修订的同时，增加了数量性状位点（QTL）、数量遗传学的新理论和新方法、进化数量遗传学、计算机在数量遗传学中的应用以及基因分子标记等全新的内容。其中，数量性状位点是目前数量遗传学与分子遗传学相结合的最新成果，连同基因分子标记，已成为当前数量遗传学中研究的热点问题。作者紧紧跟踪遗传学的发展，吸收分子生物学的最新成果，使数量遗传学以新的面貌展示在读者面前。

理论与实践相结合是本书的又一个特点。尽管本书的重点是讨论数量遗传学理论，但同时也注重数量遗传学理论在动植物改良中的应用。该书作者在写作过程中引用了大量新的应用成果和典型实验实例，做到了理论与实践相结合，引导读者把数量遗传学理论应用于动植物改良实践，指导育种工作，从而获得最佳的育种效果。翻译出版经典的数量遗传学理论著作，借鉴国外在该领域的理论和应用成果，对于在我国开展动植物改良具有重要意义。

Falconer 教授自 20 世纪 50 年代初便开始发表数量遗传学研究论文，说明他从事数量遗传学的研究工作迄今已有半个世纪，是数量遗传学领域颇有影响的学者之一，为数量遗传学的发展作出了很大贡献。他的新版《数量遗传学导论》中译本的出版，将有利于我国遗传育种界借鉴和参考。

冯志杰

2000 年 1 月于北京

第三版序言

本书写作目的是提供一本导论性的教科书，重点放在一般原理而不是实际应用上。作者努力使本书尽可能适用于更多的读者，尤其是生物学家，他们和我本人一样都只具备普通范围内的数学才能。所用数学不超出简单的代数，未列入微积分和矩阵方法，但采用了一些统计学知识，特别是方差分析以及相关和回归等方面的知识。

保持了与第二版相同的结构，但是由于包括了20年来的新发展并且植物方面的内容有所增加，故全书略有扩展。因此，对只把数量遗传学当成普通遗传学课程的部分内容的读者而言，本书包含了更为丰富的素材。而且，章节标题也方便了对有关内容的选择。现在的主要遗憾是未能更多地引用本书初版以来那些用以说明主题的实验研究。

新版的修订未增加太多的篇幅，对新进展不做详细论述，更多地是指明文献来源，但对第12章做了较大修订，因为已证明突变对数量遗传学是不可忽视的。同时对第15章和第20章作出较多的修订。

本书的习题以前都是单独出版，现在和正文放在一起，紧跟在它们所属的各章之后。这些习题有难有易，我想，对于所有学生来说，一些习题能立即解答，而另一些则需绞尽脑汁去完成。有些习题还基于前面习题的数据和答案。因此，建议学生们保留他们的计算结果供以后使用；这将避免重复计算。我尽力使习题以真实数据为基础，从而使其更富于趣味性和现实性。这样一来，计算很少能简单地解答出来，大多数习题需要一个袖珍计算器来解答。有些习题在这一版中已作修订。答案列在本书的最后，顺序安排与习题不同，以避免无意中看到下一个习题的答案。答案不仅仅是习题的最终结果，而且给出了怎样解答这些习题的较为全面的解释。

毫不夸张地说，没有Alan Robertson教授的帮助，本书不可能最初写成。我对这个学科的理解得益于同他的经常性讨论。我同样要感谢W.G. Hill教授在第二版和本版的准备过程中所给予的指导。没有他的建议，很多修订不可能做到。R.C. Roberts博士阅读了第一版和第二版的手稿，他的建议使本书作出很多改进。Paul M.Sharp博士核对了所有习题的解答并提出了很多有价值的建议。作者还得到很多其他同事的帮助，他们曾对某些问题提出建议。我深深感谢他们以及我的妻子，他们在很多方面给作者以帮助。如果书中还存在错误和疏漏的话，那完全是自己所为。对所有建议当致诚挚谢意。

Department of Genetics
West Mains Road
Edinburgh, EH9 3JN
Scotland

D.S.Falconer
1988年2月

第四版序言

数量遗传学现在正与分子遗传学相结合，与前一版相比需要更多考虑本学科中这一非常活跃的领域 因此，增加了新的一章，它是关于数量性状位点（QTL）的——引起数量变异的基因的定位和特性描述。关于自然选择的第20章，作了大量的重写，对突变和遗传变异的保持作了更全面的论述；我们希望这些增加将使本书对进化数量遗传学的学生们更有用。在较前各章中，扩展了多态性和中性突变的论述，有关近交的各章中一些部分压缩了。

我们对 James D.Fry 博士、W.G.Hill 教授、Peter D.Keightley 博士、Mark Kirkpatrick 博士和 Michael Turelli 博士的帮助表示衷心的感谢。Richard Lyman 博士制作了图21-3和图21-4，Hartwig H. Geiger 博士指出了等式［15.8］中的一个错误，现已改正过来。对二位的热情帮助，作者表示深深谢忱。最后，第一作者非常感谢 Hill 教授实验室提供的友好合作和支持。

D.S. Falconer
T.F.C. Mackay
1995 年 3 月

译者的话

让我们回顾一下《数量遗传学导论》的出版发行情况。本书前三版由 D. S. Falconer 教授独著。第一版由 Oliver & Boyd 于 1960 年出版，到 1972 年已第 5 次重印，1975 年和 1976 年又由英国 Longman Group Limited 两次重印。我国著名学者杨纪珂和汪安琦翻译了本书第一版并由科学出版社于 1965 年出版。1981 年由英国 Longman Group Limited 重新出版第二版，1982 年和 1983 年（修正了错误）两次重印，1985 年两次重印，1986 年由英国 Longman Scientific & Technical 重印。1989 年再次由英国 Longman Group Limited 出版第三版，1990 年和 1993 年两次重印。1996 年又由英国 Longman Group Limited 出版第四版，第四版由英国爱丁堡大学 D. S. Falconer 教授和美国北卡罗来纳州立大学 T. F. C. Mackay 教授合著。

本书的翻译工作实际上始于 1993 年底，我那时在攻读博士学位，利用课余时间翻译本书第三版。正当我快译完时，本书第四版又于 1996 年面世了。本书第四版新增的内容有：①数量遗传学中新理论和新方法及其在动物和植物育种以及人类疾病中的潜在应用，②阐明了对理解和使用用于数量性状分析的计算机程序必要的概念，③进化数量遗传学的内容，④数量性状位点分子标记的识别和应用，⑤第四版有参考文献 468 篇，比第三版新增 122 篇。为了对读者负责，把最新科技介绍到国内，我决定继续翻译本书第四版。1996 年 7 月 5 日我毕业分配到中国农业科学院畜牧研究所工作，在工作之余仍挤出时间坚持翻译。现在，本书第四版中译本就要由中国农业科技出版社出版发行了，希望数量遗传学领域内的这部经典著作在我国的翻译出版能对我国相关领域的科研、教学和生产起到促进作用。

我非常感谢本书作者 D. S. Falconer 教授和 T. F. C. Mackay 教授无偿提供中文版权，也感谢本书版权所有者英国 Longman Group Limited 授权中国农业科技出版社出版该书中译本在中国大陆发行。本书第四版中译本之所以能与读者见面，还在于得到中国农业科学院、中国农业科学院畜牧研究所和中国农业科技出版社领导的大力支持。中国科学院院士、中国农业大学吴常信教授以及中国农业科技出版社冯志杰副总编为这个中译本撰写了序言，中国农业大学师守堃教授和辽宁省农业科学院大连生物技术研究所王富德研究员审校了全部译文。对所有这些支持与鼓励，我谨表示衷心感谢！

由于译者水平所限，书中定有不妥和错误之处，深望同行专家和广大读者不吝赐教！

储明星

2000 年 1 月于北京

鸣　谢

对下列组织和个人允许我们复制其版权材料表示感谢：

美国数学协会，图 4-2 引自 S. Wright 的论文：'Statistical genetics and evolution' pp. 223 ~ 246 *Bulletin of the American Mathematical Society* (1942) Vol. 48；剑桥大学出版社，图 4-1 引自 S. Wright 的论文：'The genetic structure of populations' pp. 323 ~ 354 *Annals of Human Genetics* (1951) Vol. 15；J. W. Dudley 教授，图 12-3(a)引自他的论文：'76 generations of selection for oil and protein percentage in maize' in *Proceedings of the International Conference on Quantitative Genetics*: August 16 ~ 21, 1976 by E. Pollak, O. Kempthorne and T. Bailey Jr © 1977 by Iowa State University Press；美国遗传学协会，图 15-2 引自 A. Robertson 的论文：'The effect of inbreeding on the variation due to recessive genes' pp. 189 ~ 207 *Genetics* (1952) Vol. 37，图 17-2 引自 J. W. MacArthur 的论文：'Selection for small and large body size in the house mouse' pp. 194 ~ 209 *Genetics* (1949) Vol. 34，图 21-3 引自 A. E. Shrimpton and A. Robertson 的论文：'The isolation of polygenic factors controlling bristle score in *Drosophila melanogaster*. Ⅱ. Distribution of third chromosome effects with chromosome sections' pp. 445 ~ 459 *Genetics* (1988) Vol. 118，图 21-4 引自 M. C. de Vincente and S. D. Tanksley 的论文：'QTL analysis of transgressive segregation in an interspecific tomato cross' pp. 585 ~ 596 *Genetics* (1993) Vol. 134；生物科学国际联合会，图 2-2 和 2-3 引自 D. S. Falconer 的论文：'Asymmetrical responses in selection experiments' pp. 16 ~ 41 *Symposium on Genetics of Population Structure* Series B No. 15；Pergamon Journals Ltd，图 15-1 引自 M. Rasmuson 的论文：'Variation in the bristle number of *Drosophila melanogaster*' *Acta Zoolog*. No. 33；A. Robertson 教授和 Cold Spring Harbor Laboratory，图 12-3(c)引自论文：'Selection in animals: synthesis' pp. 225 ~ 229 *Cold Spring Harbor Symposium on Quantitative Biology* Vol. 20；芝加哥大学出版社，图 2-4 引自 B. Wallace 的论文：'The elimination of an autosomal lethal from an experimental population of *Drosophila melanogaster*' pp. 65 ~ 66 *American Naturalist* No. 97 (1963)；B. H. Yoo 和剑桥大学出版社，图 12-3(b)引自论文：'Long - term selection for a quantitative character in large replicate populations of *Drosophila melanogaster*' pp. 1 ~ 17 *Genetical Research* No. 35.

虽然我们尽一切努力寻找版权材料的持有者，但对少数材料并未做到，我们尽一切可能向诸位版权持有者表示歉意，我们并非故意侵犯版权持有者的权利。

绪　论

数量遗传学所涉及的是那些个体间差异的遗传性，这些差异是程度上的而不是种类上的，是数量的而不是质量的。正如达尔文所说，这些个体差异“用人类对家畜生产中一定方向的个体差异积累的同样方式，为自然选择提供可作用和可积累的材料”。因此，了解和认识这些差异的遗传特性对于进化研究以及遗传学在动植物育种方面的应用具有重要意义。正是这两个领域的需要，使数量遗传学才获得发展的主要动力。

事实上，任何物种的每一个器官和机能都表现出这种性质的个体差异，人类本身或家畜个体之间的大小差异就是大家熟悉的一个例子。个体大小从一个极端到另一个极端的一个连续等级系列，并不自然地归入界限分明的类型中。与此相反，质量差异将个体分为截然不同的类型，很少或根本没有中间过渡类型。例如，蓝眼和棕眼个体之间、血型之间、正常颜色和白化个体之间的差异就是如此。大家熟悉的反映遗传机制的孟德尔比率，只有当一单个位点上的一个基因差异引起该生物体某种特性产生很容易检测的差异时才能表现出来。数量差异，就其遗传而言，取决于微效基因，它们的效应较变异的其他成因的为小。而且，数量差异通常虽然不总是受很多位点上的基因差异的影响。因此，单个基因无论是很少的几个或者很多，都不能通过它们的分离来识别，并且也不显示孟德尔比率，因而不能应用孟德尔分析方法。

然而，数量遗传学的一个基本前提就是数量差异的遗传取决于那些与表现质量差异的基因具有相同的传递法则和相同的一般特性的基因。因此，数量遗传学是孟德尔遗传学的拓展，直接以孟德尔原理为基础。

数量遗传学的研究方法在两个方面不同于孟德尔遗传学所使用的方法。第一，既然不能观察到比率，单个后裔就不能提供足够的信息，于是研究的单元必须扩展到“群体”，也就是说，由很多后裔组成的较大的个体群。第二，要研究的数量差异的性质需要对个体进行度量，不仅仅是分类。因此，将孟德尔遗传学扩展到数量遗传学可能分两步进行，第一是引入与“群体”遗传特性有关的新概念，第二是引入与度量的遗传性有关的概念。本书就是这样进行的。第一部分，包括第 1 章到第 5 章，借助于容易引起可识别（因而是质量）差异的基因来描述群体的遗传特性。数量差异直到第二部分才讨论，第二部分从第 6 章开始，两个部分经常通过不同的名称来区分，第一部分被称为“群体遗传学”，第二部分被称为“数量遗传学”或“统计遗传学”。

数量遗传学的理论基础是在 1920 年前后通过 Fisher（1918）、Haldane（1932 年总结）和 Wright（1921）的工作而奠定的。在以后的年代中，数量遗传学的发展主要是他们和许多其他遗传学家、统计学家的详尽阐述、解释以及补充，发展成了今天我们所见到的一个为大多数人所接受且充实有效的理论体系。

当将遗传学扩展到群体特性和多位点上的基因同时分离时，其理论便由孟德尔遗传结果的推导组成。进行推导所依据的前提是数量差异的遗传借助基因，这些基因服从孟德尔的传递法则并可能具有孟德尔遗传学已知的任何特性。“可变表现”的特性被认为极其重

要并且可能成为另一个前提：表型中基因型的表达会被非遗传原因所修饰。显性、上位、基因多效性、连锁和突变特性的后果也需加以考虑。如果基因具有这些特性，那么数量遗传学理论可使我们推导一个群体的遗传特性。还可使我们预测任何特定育种计划的结果，包括自然选择的后果。因此，数量遗传学理论成为理解进化的基础。该理论的主要实际用途在于比较各种动植物改良方法的价值。

作为对理论方面的补充，数量遗传学有三个实验方面的作用。第一，群体的实验研究有利于我们推导与数量变异有关的基因特性。第二，实验育种有利于我们验证理论的有效性。第三，有一些育种方法的结果不能依理论来预测，有关这些方法的问题只有通过实验才能回答。现在有大量的实验资料比较详细地证实了这一理论，表明与数量变异有关的基因确实具有孟德尔遗传学特性，而且在一定的置信度上能够预测大多数育种方法的结果。目的是要描述所有已经合理稳固建立起来的材料，并且为清晰起见，在不引起误解的情况下尽可能地加以简化。因此，我们重点放在理论方面。虽然从实验资料中经常可以直接得出结论，但这个学科的实验方面主要以实例的形式出现，以例证理论上的结论为目的来选择例子。但是并非所有这些例子都能证明它们要例证结论的有关假设。实验结果常常可以有多种解释。所提及的实验工作只是已做实验工作的一个很小的并且往往也不是随机的样本。特别在植物和家畜方面，已有的实验工作比本书所引用的要多得多。

作者并不想在这个学科的任何部分把所有发表过的文献都加以罗列；也不试图标明那些观点来源或追溯其历史。假使这样做，则需要写一本厚得多的书，也就不能做到言简意赅。本书大部分取材于下面列出的一些著作，但并非经常加以引用。当表述任何一个结论而对它的推导没有给出充分的解释时，则在书中列出文献出处。这些参考文献虽非总是原始的论文，但均是较新的文献，读者从中将易于找到所讨论的主题。其中包括一部对数量遗传学发展产生过重大影响的论文集《遗传学论丛》（第 15 卷），文集中附有 Hill（1984）的详尽评述。

主要参考书目

（全部参考书目见所列“参考文献”）

Becker（1984）*Manual of Quantitative Genetics*.

Bulmer（1985）*The Mathematical Theory of Quantitative Genetics*.

Crow（1986）*Basic Concepts in Population, Quantitative, and Evolutionary Genetics*.

Crow and Kimura（1970）*An Introduction to Population Genetics Theory*.

Hartl and Clark（1989）*Principles of Population Genetics*.

Hedrick（1985）*Genetics of Populations*.

Jacquard（1974）*The Genetic Structure of Populations*.

Kempthorne（1957）*An Introduction to Genetic Statistics*.

Kimura（1983）*The Neutral Theory of Molecular Evolution*.

Li（1976）*First Course in Population Genetics*.

Mather and Jinks（1977）*Introduction to Biometrical Genetics*.（1982）*Biometrical Genetics*.

Mayo（1987）*The Theory of Plant Breeding*.

Weir（1990）*Genetic Data Analysis*.

Wright（1968～1978）*Evolution and the Genetics of Populations*, Vols 1～4.

目　　录

中文版序言（一）
中文版序言（二）
译者的话
第四版序言
第三版序言
鸣　谢
绪　论

第 1 章　群体的遗传组成 …… (1)
基因和基因型的频率 …… (1)
自然群体中的孟德尔变异 …… (3)
变化的原因 …… (3)
哈迪—温伯格平衡 …… (4)
哈迪—温伯格法则 …… (4)
哈迪—温伯格法则的应用 …… (6)
交配频率和哈迪—温伯格法则的另一种证明 …… (9)
复等位基因 …… (10)
性连锁基因 …… (11)
多于一个位点 …… (13)
非随机交配 …… (16)
选型交配 …… (16)
习　题 …… (17)

第 2 章　基因频率的变化 …… (19)
迁　移 …… (19)
突　变 …… (19)
非频发突变 …… (19)
频发突变 …… (20)
选　择 …… (21)
选择下基因频率的变化 …… (22)
选择的有效性 …… (24)
需要的世代数 …… (26)
平均适合度和负荷 …… (27)

平 衡 …… (27)
突变和选择之间的平衡 …… (27)
平衡的改变 …… (29)
有利于杂合体的选择 …… (30)
多态性 …… (33)
习 题 …… (36)

第3章 小群体：Ⅰ. 简化条件下基因频率的变化 …… (38)
理想群体 …… (39)
抽样 …… (40)
基因频率的方差 …… (40)
固 定 …… (43)
基因型频率 …… (44)
近 交 …… (45)
理想群体中的近交 …… (46)
基因频率的方差 …… (48)
基因型频率 …… (48)
习 题 …… (50)

第4章 小群体：Ⅱ. 较少简化条件 …… (52)
有效群体大小 …… (52)
极近亲交配的排除 …… (52)
雄性和雌性数目不同 …… (53)
连续世代中数目不等 …… (53)
家系大小的非随机分布 …… (54)
最低程度的近交 …… (55)
世代重叠 …… (55)
突变、迁移和选择 …… (57)
非频发中性突变 …… (57)
频发突变和迁移 …… (58)
选 择 …… (60)
自然群体中的随机漂变 …… (60)
多态性 …… (61)
中性理论 …… (62)
习 题 …… (63)

第5章 小群体：Ⅲ. 系谱群体和近亲近交 …… (65)
系谱群体 …… (65)
一个个体的近交系数 …… (65)

共祖率或亲缘关系 …… (67)
规则的近交系统 …… (69)
近亲近交 …… (70)
固 定 …… (71)
反复回交 …… (72)
杂交和后续世代 …… (73)
混合的近交和杂交 …… (74)
基础的改变：结构化群体 …… (75)
突 变 …… (76)
有利于杂合体的选择 …… (76)
习 题 …… (77)

第6章 连续变异 …… (79)
数量性状 …… (80)
数量性状的特性 …… (82)
习 题 …… (84)

第7章 数值与均数 …… (86)
群体平均值 …… (87)
平均效应 …… (89)
育种值 …… (91)
显性离差 …… (92)
互作离差 …… (94)
习 题 …… (95)

第8章 方 差 …… (97)
方差组分 …… (97)
作为总方差部分的组分 …… (97)
遗传决定程度 V_G/V_P 的估计 …… (98)
方差的遗传组分 …… (99)
加性和显性方差 …… (99)
总遗传方差 …… (100)
互作方差 …… (103)
由不平衡引起的方差 …… (103)
基因型和环境之间的相关与互作 …… (104)
相 关 …… (104)
互 作 …… (104)
环境方差 …… (106)
多次度量：重复力 …… (108)

方差分剖的总结 …… (113)
习 题 …… (113)

第 9 章 亲属间的相似性 …… (115)
遗传协方差 …… (116)
后代与一个亲本 …… (116)
后代与中亲值 …… (117)
半同胞 …… (118)
全同胞 …… (120)
孪生子 …… (121)
概 要 …… (121)
上位互作 …… (122)
环境协方差 …… (123)
表型相似性 …… (125)
习 题 …… (125)

第 10 章 遗传力 …… (127)
遗传力的估计 …… (129)
亲子回归 …… (130)
同胞分析 …… (131)
后代对母畜的公畜内回归 …… (134)
联合估计值 …… (134)
双胞胎和人类数据 …… (135)
选型交配 …… (137)
估计值的精确度和实验设计 …… (140)
亲子回归 …… (140)
同胞分析 …… (142)
亲本选择 …… (142)
习 题 …… (143)

第 11 章 选择：Ⅰ．反应及其预测 …… (145)
选择反应 …… (146)
反应的预测 …… (147)
选择差和选择强度 …… (148)
反应的改良 …… (150)
反应的度量 …… (152)
世代平均值的变异性 …… (152)
加权选择差 …… (154)
现实遗传力 …… (155)

母体效应 …… (156)
人工选择下基因频率的变化 …… (156)
选择对方差的影响 …… (158)
习 题 …… (161)

第 12 章 选择：Ⅱ．实验结果 …… (164)
短期结果 …… (164)
反应的重复性 …… (164)
抽样方差 …… (165)
反应的不对称性 …… (166)
长期结果 …… (169)
选择极限 …… (169)
突 变 …… (173)
选择极限的原因 …… (175)
位点（有效因子）数目和标准化效应 …… (176)
习 题 …… (177)

第 13 章 选择：Ⅲ．来自亲属的信息 …… (179)
选择标准 …… (180)
简单方法 …… (180)
反应的预测 …… (182)
合并选择 …… (184)
各种方法的相对优点 …… (186)
指数选择 …… (188)
指数的构建 …… (188)
准确度 …… (190)
选择反应 …… (191)
实际成就 …… (192)
习 题 …… (192)

第 14 章 近交和杂交：Ⅰ．平均值的变化 …… (194)
近交衰退 …… (194)
选择的效应 …… (197)
杂种优势 …… (199)
单 交 …… (200)
习 题 …… (205)

第 15 章 近交和杂交：Ⅱ．方差的改变 …… (207)
近 交 …… (207)
遗传方差的再分配 …… (207)
环境方差 …… (210)
近交系的一致性 …… (211)
突 变 …… (212)
亚系歧化 …… (213)
杂 交 …… (214)
杂种之间的方差 …… (214)
配合力 …… (215)
习 题 …… (219)

第 16 章 近交和杂交：Ⅲ．应用 …… (222)
配合力的选择 …… (223)
三元和四元杂交；回交 …… (224)
正反交反复选择（RRS） …… (226)
超显性 …… (226)
自然自体受精植物 …… (227)
习 题 …… (227)

第 17 章 尺 度 …… (229)
分布与方差 …… (229)
互 作 …… (232)
结 论 …… (234)
习 题 …… (234)

第 18 章 阈性状 …… (236)
趋向和阈值 …… (236)
两种类别，一个阈值 …… (236)
趋向模型的充分性 …… (239)
尺度关系 …… (240)
三种类别，两个阈值 …… (240)
阈性状的选择 …… (243)
习 题 …… (245)

第 19 章 相关性状 …… (247)
遗传和环境相关 …… (247)
遗传相关的估计 …… (249)

选择的相关反应 …… (250)
间接选择 …… (252)
基因型—环境互作 …… (254)
指数选择 …… (257)
指数的构建 …… (257)
反 应 …… (259)
选择对遗传相关的影响 …… (262)
习 题 …… (263)

第 20 章 自然选择下的数量性状 …… (265)
自然选择 …… (265)
适合度及其组分 …… (265)
数量性状和适合度之间的关系 …… (267)
"适合度轮廓" …… (267)
对自然选择的反应 …… (268)
适合度 …… (268)
相关反应 …… (268)
选择的强度 …… (269)
平衡群体 …… (270)
适合度 …… (270)
主要组分 …… (271)
具有中间值最优的性状 …… (272)
具有中间值最小适合度的性状 …… (274)
中性性状 …… (274)
突变引起的变异来源 …… (275)
突变方差 …… (275)
遗传变异的保持 …… (276)
中性突变和随机漂变之间的平衡 …… (277)
突变—选择平衡 …… (277)
习 题 …… (279)

第 21 章 数量性状位点 …… (281)
主效基因 …… (281)
检测方法 …… (282)
QTL 作图的方法 …… (283)
标记位点 …… (284)
QTL 基因型 …… (284)
单个标记分析 …… (285)
区间作图分析 …… (286)

遗传和统计上的考虑 …… (288)
实验设计 …… (288)
多重测验 …… (289)
最大似然估计 …… (289)
多个 QTL …… (290)
实验结果 …… (291)
位点数目 …… (291)
基因效应 …… (292)
一致性 …… (294)
从 QTL 到基因 …… (294)
习 题 …… (296)

附录表 …… (298)
符号集注 …… (300)
Mather 和 Jinks 使用的符号的等价意义 …… (303)
习题解答 …… (304)
参考文献 …… (351)

第 1 章　群体的遗传组成

基因和基因型的频率

描述一群生物个体的遗传组成，我们应详细说明它们的基因型，并说出每一种基因型有多少。假若我们不在意基因型间表型差异的性质，那么这将是一个完整的描述。为简单起见，假设我们关注某一个常染色体位点 A，并且这个位点上的两个不同的等位基因 A_1 和 A_2 存在于个体之间，于是将有三种可能的基因型 A_1A_1、A_1A_2 和 A_2A_2（这里我们只关注二倍体的生物体，全书均如此）。群体的遗传组成可通过属于每一种基因型的个体比例或百分率，或换句话说，通过个体之间三种基因型的频率进行充分的描述。这些比例或频率被称为基因型频率，一种特定基因型的频率就是它在个体之间的比例或百分率。例如，如果我们发现群体中四分之一的个体是 A_1A_1 型的，则这种基因型的频率就是 0.25 或 25%。当然，所有基因型的频率加在一起一定等于 1 或 100%。

例 1.1

人类 M-N 血型是由一个位点上两个等位基因决定的，三种基因型对应三种血型 M、MN 和 N。以下的数据取自 Mourant（1954）的资料（表 1-1），显示了东格陵兰岛爱斯基摩人和冰岛人的血型分布频率。

表 1-1　东格陵兰岛爱斯基摩人和冰岛人的血型分布频率

		血　型			个体数目
		M	MN	N	
频率（%）	格陵兰人	83.5	15.6	0.9	569
	冰岛人	31.2	51.5	17.3	747

很明显，这两个群体在这些基因型频率上不同，N 血型在格陵兰人中稀少而在冰岛人中相对普遍。这个位点不仅是两个群体各自变异的一种来源，而且也是群体间遗传差异的一种来源。

在遗传学意义上，一个群体不仅仅是一群个体，而且是一个能繁殖的类群；一个群体的遗传学不仅涉及个体的遗传组成，而且也涉及从一个世代到下一世代的基因传递。在传递过程中，亲本的基因型分开了，后裔中由配子传递的基因组成了一组新的基因型。这样，群体携带的基因在世代间具有连续性，但它们构成的基因型却没有这种连续性。一个群体的遗传组成，以它携带的基因为依据，可通过基因频率的阵列来描述；也就是说，通过在每一个位点上存在的等位基因以及在每一个位点上不同等位基因的数目或比例加以详细说明。例如，如果 A_1 是位点 A 上的一个等位基因，那么 A_1 基因的频率，或 A_1 的基因频率是这个位点上所有基因中 A_1 等位基因所占的比例或百分率。任何一个位点上所有等

位基因的频率加起来一定等于 1 或 100%。

一组个体之间某特定位点上的基因频率可根据已知的基因型频率来确定。举一个假想的例子，假定有两个等位基因 A_1 和 A_2，我们对 100 个个体进行分类并计数每一种基因型的数目（表 1-2）。

表 1-2　根据基因型频率确定基因频率

		A_1A_1	A_1A_2	A_2A_2	总数	
个体数		30	60	10	100	
基因数	A_1	60	60	0	120	200
	A_2	0	60	20	80	

每一个体包含两个基因，这样我们计数了这个位点上的 200 个基因的代表。每一个 A_1A_1 个体包含两个 A_1 基因，每一个 A_1A_2 个体包含一个 A_1 基因。因此，样本中有 120 个 A_1 基因和 80 个 A_2 基因。进而 A_1 的频率是 60%或 0.6，A_2 的频率是 40%或 0.4。为了用一种更一般的形式表示这种关系，可以设定基因和基因型的频率（表 1-3）。

表 1-3　基因频率和基因型频率的关系

	基因		基因型		
	A_1	A_2	A_1A_1	A_1A_2	A_2A_2
频率	p	q	P	H	Q

这样，$p+q=1$ 以及 $P+H+Q=1$。因为每一个体包含两个基因，A_1 基因的频率为 $\frac{1}{2}(2P+H)$，而计数过的个体间基因频率和基因型频率之间的关系如下：

$$\left.\begin{aligned} p &= P+\frac{1}{2}H \\ q &= Q+\frac{1}{2}H \end{aligned}\right\} \qquad [1.1]$$

例 1.2

为了说明从基因型频率中计算基因频率，我们可以取例 1.1 给出的 M-N 血型频率。M 和 N 血型表示两个纯合基因型，MN 血型表示杂合基因型。根据等式 [1.1]，格陵兰人中 M 基因的频率是 $0.835+\frac{1}{2}(0.156)=0.913$，N 基因的频率是 $0.009+\frac{1}{2}(0.156)=0.087$，频率总和等于 1.000。

表 1-4　格陵兰人和冰岛人血型基因的频率

	基因频率（%）	
	M	N
格陵兰人	91.3	8.7
冰岛人	57.0	43.0

对冰岛人样本进行同样的运算，我们发现两个群体中有如表 1-4 所示的基因频率（以百分率表示）。这样，两个群体在基因频率和基因型频率上都不同。

自然群体中的孟德尔变异

我们对自然群体中作为孟德尔单位分离的不连续性状能够观察到很多不同水平的遗传变异。对表型（例如植物花色、蜗牛壳色和图案）具有巨大影响的可见变异或者像侏儒这样的巨大突变都在变异系谱的一端。但是，个体之间的很多变异在表型上并不产生明显的差异。这种隐蔽的变异已被研究蛋白质以及 DNA 本身差异的方法所揭示。例 1.1 和例 1.2 中的 MN 血型变异例释了一种通过抗体反应可检测出来的隐蔽变异。在电场存在的情况下，蛋白质电泳是检测可溶性蛋白质在凝胶上流动性差异的一种技术。这种方式检测出的变异是作为共显性等位基因遗传的，被称为等位酶。它们是由引起蛋白质电荷改变的氨基酸替代所造成的。因此，蛋白质电泳可检测蛋白质之间大约 25% 的氨基酸差异，因为 20 种氨基酸有 5 种是带电荷的。

采用限制性酶可检测 DNA 序列的变异。这些酶识别特异性的 4 或 6 个碱基的 DNA 序列，这些序列无论何时出现，这些酶都可以切割它。然后，这些 DNA 片段通过电泳按大小被分开，并通过称为 Southern blotting 的程序与一种标记的 DNA 探针杂交而被观察。在探针可识别的 DNA 区间内，如果个体之间在限制性位点上存在变异，那么这种变异会被斑点内限制性片段的大小变化所揭示，这被称为限制性片段长度多态性，或 RFLPs。与只能检测功能蛋白质变化的蛋白质电泳不同，RFLP 变异可以位于基因组的非编码和编码区域。虽然克隆 DNA 对检测这种变异是必要的，但不必知道克隆探针的功能或染色体位置。限制性片段长度的变异还可由两个限制性位点之间的 DNA 序列的插入和缺失引起。大的插入通常是转座（位）因子：它的 DNA 序列在基因组中是以多个、分散的拷贝存在的，能够从一个位置移动到另一个位置。另外的长度变异是由“小卫星”或“微卫星”位点上串联重复 DNA 序列的数目变化所引起。前者，也称为 VNTR（可变数目的串联重复）位点，由 10～60 个碱基对长的重复单位组成。微卫星（或简单序列重复，SSR）位点由 1～6 个碱基对这样较短的重复单位组成，例如，$(CA)_n$ 或 $(AGC)_n$，其中 n 是重复单位数目，是可变的。最后，通过比较由直接测序所得的实际 DNA 序列来最终分辨个体间的变异。

对不连续性状的等位变异，不管是表型上可见的还是隐蔽的，都被称为多态性，关于这些，将在第 2 章和第 4 章进行更多的说明。多态位点导致数量性状的变异，这是本书的主题。

变化的原因

在基因从一个世代传递到下一世代过程中，几种作用影响基因和基因型频率。为了全面理解数量遗传变异，我们需要了解这些因素如何单独和共同影响群体随时间的遗传变异，以及作为基因频率改变的作用，它们的相对重要性有多大。这些作用形成了下面四章的主要内容，在此我们对这些作用做简要评述，以便说明本章对什么因素不予考虑的想法。可引起群体遗传特性变化的作用是：

群体大小　从一个世代传递到下一世代的基因是亲代基因的一个样本。因此，连续世代间的基因频率存在着抽样变异，而且亲本数量越小，抽样变异就越大。抽样变异的效应子第 3～5 章中给予考虑，但总是假定我们在处理一个“大群体”，这只意味着这个群体中的抽样变异小到可以忽略不计，因此，届时讨论中不涉及它。为实用目的，一个“大群

体”是成年个体数以百计而不是数以十计的一个群体。

生殖力和生活力的差异 虽然现在我们不涉及所讨论的基因表型效应，但我们不能忽视它们对生殖力和生活力的影响，因为这些效应影响下一代的遗传组成。亲本之间不同的基因型可能有不同的生殖力，如果是这样，则它们对形成下一世代用的配子贡献也不会相等。通过这种方式，基因频率在传递过程中可能会改变。而且，新形成的合子之间的基因型可能有不同的存活率，因此到个体成熟且它们本身成为亲本的时候，新的世代中基因频率也可能会改变。这些过程被称为选择，将在第 2 章描述。其时我们将假定这些过程不起作用。我们可以取人类血型基因来说明这一目的，因为对这些基因起作用的选择力量可能不强烈。相反地，产生一种与野生型相比是异常突变表型的基因，经常承受着更多的选择。

迁移和突变 群体的基因频率还可能因另一个群体中的个体迁入以及基因突变而改变。这些过程将在第 2 章描述，在这里也假定它们不起作用。

交配系统 后裔的基因型是由配子成对结合所形成的合子决定的，配子的结合又受到亲本交配的影响。因此，后裔世代中基因型频率受亲本世代中交配配偶的基因型的影响。我们将首先假定，就讨论的基因型而言，交配是随机的。随机交配（random mating 或 panmixia）意味着任一个体与群体中的任何其他个体都有相等的交配机会。重要之点在于交配的个体不应该有基因型的相似、或通过祖先而彼此有相关的特殊倾向。如果一个群体占有广大的地理区域，那么居住在同一地区的个体比居住在不同地区的个体似乎有更多的交配，因此，交配的配偶倾向于通过祖先而相关。为此一个分布很广的群体很可能被细分成很多地区群，并且仅在地区群内交配是随机的。细分群体的特性依赖于地区群的大小，这将在第 3～5 章群体大小的效应中再来描述。

哈迪—温伯格平衡

哈迪—温伯格法则

在一个没有选择、突变和迁移的大随机交配群体中，基因频率和基因型频率在世代间保持不变；而且，基因频率和基因型频率之间存在一种简单的关系。一个群体的这些特性是从众所周知的哈迪—温伯格法则这一定理（或原理）导出的，哈迪和温伯格各自于 1908 年证明了这些特性。一个具有稳定的基因和基因型频率的群体被认为处于*哈迪—温伯格平衡*。基因频率和基因型频率之间的关系是最重要的，因为有关群体遗传学和数量遗传学的很多推论都基于它。关系是这样的：如果亲本之间两个等位基因的基因频率是 p 和 q，那么后裔之间的基因型频率是 p^2、$2pq$ 和 q^2，因此有：

	亲本中基因		后裔中基因型			
	A_1	A_2	A_1A_1	A_1A_2	A_2A_2	
频率	p	q	p^2	$2pq$	q^2	[1.2]

（上面的关系指的是常染色体基因；性连锁基因并不这么简单，将在以后解释。）保持哈迪—温伯格法则所需要的随机交配和没有选择的条件仅仅指的是所考察的基因型。关于其他属性可能有优先交配，其他位点的基因型可能受到选择，但不影响正在讨论的问题。

两个附加的条件是基因在配子发生过程中正常分离以及雄性和雌性中基因频率相同。这些要求的理由将在证明过程中看出来。

哈迪—温伯格法则的证明包含 4 个步骤，它们汇总在表 1-5 中，同时也给出了确保每一步推导有效所必须保持的条件。4 个步骤详述如下。

表 1-5　在证明哈迪—温伯格法则过程中推导的步骤以及必须保持的条件

步骤	推导（起：止）	条　件
1a	亲本基因频率 所有配子的基因频率	(1) 正常的基因分离 (2) 亲本具有相等的生殖力
1b	所有配子的基因频率 形成合子的配子的基因频率	(3) 配子有相等的受精能力 (4) 大群体
2	形成合子的配子的基因频率 合子的基因型频率	(5) 随机交配 (6) 雄亲和雌亲有相等的基因频率
3	合子的基因型频率 后裔中基因型频率	(7) 相等的存活力
4	后裔中基因型频率 后裔中基因频率	

1. 从亲本基因频率到配子基因频率　假设亲本世代具有如下的基因和基因型频率（表 1-6）：

表 1-6　亲本世代的基因和基因型频率

	基　因		基因型		
	A_1	A_2	A_1A_1	A_1A_2	A_2A_2
频率	P	q	P	H	Q

产生两种类型的配子，一些配子携带 A_1，另一些配子携带 A_2。A_1A_1 个体只产生 A_1 配子。正常分离的 A_1A_2 个体产生同等数目的 A_1 和 A_2 配子。于是，只要所有基因型都具有相等的生殖力，整个群体产生的所有配子之间，A_1 的频率就是 $P+\frac{1}{2}H$，由等式［1.1］可知，它是产生配子的亲本中 A_1 基因的频率。这样整个配子产出中基因频率与亲本的相同。这是表 1-5 中的步骤 1a。只有一些配子形成将成为下一世代中个体的合子。只要携带不同等位基因的配子在其受精能力上没有差异，并且只要形成的合子代表亲本配子的一个大样本，合子中基因频率就不会改变。这是步骤 1b。

2. 从配子基因频率到合子的基因型频率　个体间随机交配等于它们的配子间随机结合。于是合子（受精卵）之间的基因型频率就是结合起来以产生它们的配子类型的频率之积。因此，随机交配产生的后裔之间的基因型频率可简单地通过将每一性别亲本产生的配子类型频率相乘来确定。只要两性中配子频率相同，产生的合子就如表1-8所示。

表 1-7　合子的基因型频率

	基因型		
	A_1A_1	A_1A_2	A_2A_2
频率	p^2	$2pq$	q^2

A_1 卵子和 A_2 精子的结合不必与 A_2 卵子和 A_1 精子的结合加以区别；因此得到如表 1-7 所示的合子基因型频率。

表 1-8　由两性配子得到的合子基因型频率

		雌性配子及其频率	
		A_1	A_2
		p	q
雄性配子及其频率	A_1	A_1A_1	A_1A_2
	p	p^2	pq
	A_2	A_1A_2	A_2A_2
	q	pq	q^2

3. 从合子到成年　正如［1.2］表述的那样，上面推导的合子基因型频率是哈迪—温伯格频率。可是，这不是证明的彻底结束，因为除非合子有相等的存活力，至少直到能够区分它们的基因型时，才能观察到这些频率。这个步骤可能是不重要的，但是如果要了解不同的存活力影响，就必须认识这一步骤。

4. 后裔中从基因型频率到基因频率　这个最后步骤证明基因频率没有改变。只要后裔中不同基因型相等地存活到成年，此时它们能成为亲本，它们的频率就是上面给出的那样。于是成熟后裔中基因频率可通过等式［1.1］找到。A_1 的频率是 $p^2 + (2pq)/2 = p(p+q) = p$，它与亲代的相同。这证明基因频率在世代间保持不变。

现在可以陈述哈迪—温伯格法则两个深层次的方面。首先，因为亲本和后裔中基因频率相同，［1.2］中的基因频率和基因型频率之间的关系可适用于一单个世代。其次，后裔的基因型频率仅取决于亲本的基因频率而与亲本的基因型频率无关。这一点可以从上面步骤 1 中看出，其中配子类型的频率与亲本的基因频率相等，而不管基因型频率是多少。因而，具有任何基因型频率的亲本，只要它们随机交配并且雄亲和雌亲中的基因频率相同，就以哈迪—温伯格比例产生后裔。如果已知雄亲和雌亲中基因频率不相同，通过在步骤 2 中将适宜的配子频率代入表1-8，可很容易地推导出后裔的基因型频率。一个常染色体基因的基因频率在后裔的两个性别中是相等的，并且随机交配的下一代产生与原始基因频率平均值对应的哈迪—温伯格基因型频率。

处于哈迪—温伯格平衡的一个群体中基因频率和基因型频率之间的关系显示在图 1-1 中。基因型频率的曲线显示了两个重要的特征。首先，杂合体的频率不可能超过 50%，当基因频率为 $p = q = 0.5$ 时，这个最大值才出现。其次，当一个等位基因的频率低时，这个稀有的等位基因主要出现于杂合体中，纯合体几乎没有。这对选择的有效性具有重要的影响，这一点将在下一章看到。

哈迪—温伯格法则的应用

哈迪—温伯格法则在三个方面特别有用，现在来说明它们。

隐性等位基因的基因频率　本章开始我们在等式［1.1］中看到，一群个体之间的基因频率可如何从它们的基因型频率中来确定；但对此知道所有 3 种基因型的频率是必要的。继而，当杂合体不能与显性纯合体区分时，等式［1.1］的关系不适用于隐性等位基

因的情况。可是，如果基因型符合哈迪—温伯格的比例，我们就不必知道所有 3 种基因型的频率。例如，假定用 a 表示频率为 q 的一个隐性基因；那么 aa 纯合体的频率就是 q^2，并且基因频率是纯合体频率的平方根。例 1.3 说明了这种计算。要确保这种估计基因频率的方法有效，在计数纯合体之前没有选择淘汰显然是必需的。还应当注意到采用这种方法对基因频率进行估计时，对非随机交配的影响是相当敏感的。

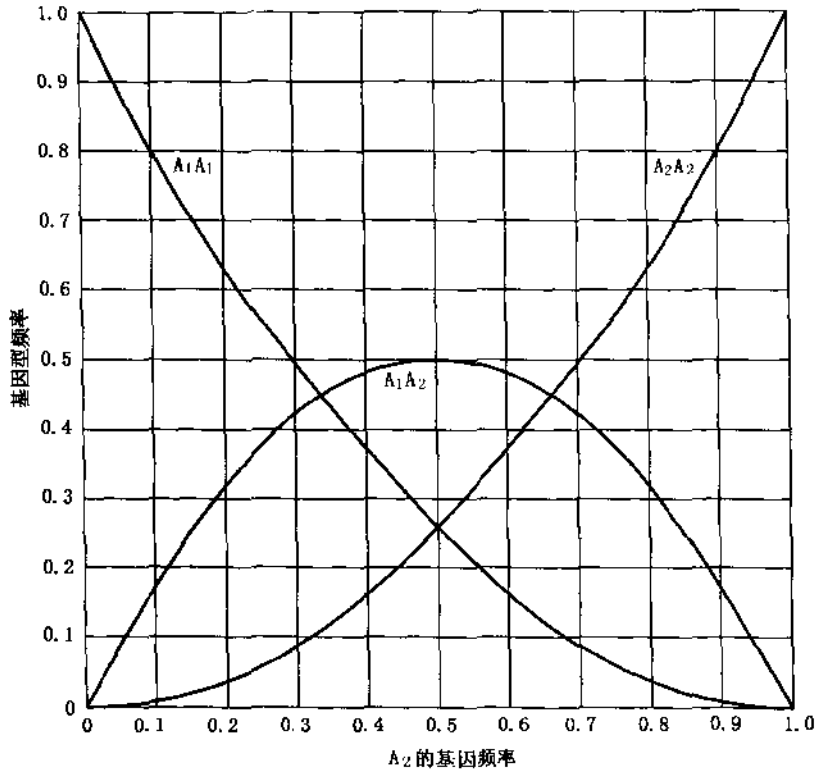


图 1-1　处于哈迪—温伯格平衡的一个群体中两个等位基因的基因型频率和基因频率之间的关系

例 1.3

苯丙酮尿症（PKU）是由单个隐性基因引起的一种人类代谢病。纯合体可在出生后几天检测出来，在那以前假定选择淘汰是可以忽略不计的。对英国伯明翰出生的婴儿经过 3 年的测验，从 55 715 个婴儿中检测出 5 例（Raine 等，1972）。样本中纯合体的频率是 90×10^{-6}或约为 1/11 000。纯合体的哈迪—温伯格频率是 q^2，因此基因频率是 $q=\sqrt{90\times10^{-6}}=9.5\times10^{-3}=0.0095$。

整个群体中杂合体的频率是$2q(1-q)$，正常个体内杂合体的频率是$2q/(1+q)$。两者计算出来的答案近似为 0.019。因此，大约正常人的 2%或每 50 人中有 1 人是 PKU 基因携带者。大多数人在发现一个稀有的隐性畸形体的杂合体竟如此的普遍时，往往会感到吃惊。作为从图 1-1 中得出的一个结论，已经注意到了这一点。

“携带者”的频率　了解隐性畸形体杂合体或“携带者”的频率经常是重要的，如果

基因频率已知就可以计算它。假定存在哈迪—温伯格平衡，则所有个体（包括纯合体）中的杂合体频率由$2q(1-q)$给出。可是，往往了解正常个体内杂合体的频率是更妥当的，虽然这在纯合体稀少时不会有很大的不同。正常个体中杂合体的频率，用符号H'表示，是基因型频率Aa/(AA + Aa)的比率，其中a是隐性等位基因。因此，当a的频率是q时，有：

$$H' = \frac{2q(1-q)}{(1-q)^2 + 2q(1-q)} = \frac{2q}{1+q} \qquad [1.3]$$

哈迪—温伯格平衡的测验　如果能识别一个位点上的所有基因型，其所得数据可用，那么可以检验基因型的观察频率是否与处于哈迪—温伯格平衡的群体相一致。根据哈迪—温伯格法则，后裔的基因型频率由它们亲本的基因频率决定。如果群体处于平衡，则亲本与后裔的基因频率相同，因此，后裔中观察到的基因频率，就像是亲本的基因频率，能用来计算依哈迪—温伯格法则期望的基因型频率。这种方法例证在例1.4中。

例1.4

例1.1给出了冰岛人M-N血型频率。样本的观察数在下表中。样本的基因频率首先通过等式［1.1］从观察数中计算。然后，通过等式［1.2］从基因频率中计算出哈迪—温伯格基因型频率p^2、$2pq$和q^2，每一个与总数相乘得到期望值。例如，MM的期望值是$(0.5696)^2 \times 747$。比较观察数和期望值表明，两种纯合体不足，杂合体过量。用χ^2测验观察数和期望值一致程度的好与坏。不一致性不显著，这在抽样过程中很容易偶然产生。注意这个χ_1^2只有一个自由度，因为从数据中已估计了基因频率，这样观察数和期望值在它们的基因频率和总数上必定一致（表1-9）。

表1-9　冰岛人M-N血型频率的χ^2测验

	基因型				基因频率	
	MM	MN	NN	总数	M	N
观察数	233	385	129	747	0.5696	0.4304
期望值	242.36	366.26	138.38	747		
$\chi_1^2 = 1.96$	$P \sim 0.2$					

对符合平衡群体的测验就是对哈迪—温伯格基因型频率产生的条件是否满足的测验。可是，从测验中能获得的结论却有限。当发现良好符合时，测验没有理由怀疑所有条件的满足。对血型基因作出的测验几乎总是显示相当良好的符合，正如例1.4所示。但是有一个条件，它的不满足并不导致不符合，这就是亲本之间相等的生殖力。其理由将马上给予解释。如果测验显示出观察和期望频率之间不一致，那么我们能得出结论：一个或更多的条件没有满足。但是不一致性不允许我们鉴别出其来源，或确定哪一个条件没满足。其理由是同样的不一致性可由不同的原因引起。例如，杂合体过多可由纯合体的选择淘汰引起，也可由亲本世代中雄性和雌性的基因频率不同引起。这种测验并不像它表现的那么简单，我们必须更仔细地观察这种测验能够做什么。

哈迪—温伯格法则将亲本的基因和后裔的基因型关联起来。因此，要全面测验它，我们需要知道亲本的基因频率并从亲本基因频率中计算出后裔的期望基因型频率。除描述的测验外，我们只有后裔。我们通过计数找出它们的基因频率。于是我们认为：如果这就是产生这些后裔的配子之间的基因频率，那么基因型应当像从观察的基因频率中计算出的那样处于哈迪—温伯格比例。如果亲本的基因频率和后裔的不同，那么我们已经使用了错误的基因频率来计算期望值。参考表 1-5 将指明测验的条件是随机交配，两性亲本具有相等的基因频率和后裔之间具有相同的生活力；但没有测验亲本之间相等的生殖力。因此，选择可通过生殖力起作用，并且不会被这种测验检出。通过对后裔生活力起作用的选择将导致观察和期望频率之间的不符合。可是，鉴定出降低生活力的基因型是不可能的。在论及选择的效应之后，下一章将解释其理由。关于测验局限性的更全面的讨论见 Wallace (1958, 1968)，Prout (1965)；其统计方面的更全面考察见 Smith (19[illegible]

交配频率和哈迪—温伯格法则的另一种证明

现在让我们更仔细地观察一个随机交配群体的繁殖结构，按照配偶的基因型来区别交配类型，以及观察不同交配类型的后裔之间的基因型频率是多少。这将为叙述连续世代中的基因型频率提供一个一般的方法，这将在下一章中用到。它还为哈迪—温伯格法则提供了另一种证明；这种证明比已经给出的更麻烦，但更清晰地显示了哈迪—温伯格频率是如何从孟德尔分离法则中产生的。这个过程首先是要根据亲本之间基因型的频率获得所有可能交配类型的频率，然后要根据孟德尔比率获得每一种交配类型的后裔之间的基因型频率。

考察具有两个等位基因的一个位点，并且像以前一样，假定亲本中的基因和基因型频率如表 1-10 所示：

表 1-10　亲本的基因和基因型频率

	基　因		基因型		
	A_1	A_2	A_1A_1	A_1A_2	A_2A_2
频率	p	q	P	H	Q

总共有 9 种交配类型，当交配是随机时，它们的频率可通过表1-11下角显示的频率相乘而获得。因为亲本的性别对此无关，一些交配类型是等价的，所以不同类型的数目简化成 6 种。通过对等价类型频率的求和，我们获得表1-12前两栏中所示的交配类型的频率。现在我们必须考察由每一种交配类型产生的后裔的基因型，并找出在整个后裔中每一种基因型的频率，当然得假定所有的交配类型具有相等的生殖力以及所有的基因型具有相等的生活力。这列在表 1-12 的右边栏中。这样，例如交配类型 $A_1A_1 \times A_1A_1$ 仅产生 A_1A_1 后裔。因此，整个后裔中，P^2 部分是从这种交配类型中获得的 A_1A_1 基因型。类似地，$A_1A_2 \times A_1A_2$ 交配的后裔中有 $\frac{1}{4}$ 是 A_1A_1。因此，这种频率为 H^2 的交配类型贡献给总的 A_1A_1 后裔部分是 $\frac{1}{4}H^2$。为了求出整个后裔中每一种基因型的频率，我们将每种交配类型贡献的频率相

加。简化之后的总和列在表1-12的底部，从等式［1.1］给出的恒等式可知，它们等于 p^2、$2pq$ 和 q^2。这些就是哈迪—温伯格平衡频率，并且我们已经证明它们可以通过一代随机交配而达到，不管亲本之间的基因型频率如何。

表 1-11 两个等位基因随机交配时各种基因型的频率

			雄性亲本的基因型及其频率		
			A_1A_1	A_1A_2	A_2A_2
			P	H	Q
雄性亲本的基因型及其频率	A_1A_1	P	P^2	PH	PQ
	A_1A_2	H	PH	H^2	HQ
	A_2A_2	Q	PQ	HQ	Q^2

表 1-12 两个等位基因随机交配基因型频率的计算

交配		后裔的基因型及其频率		
类型	频率	A_1A_1	A_1A_2	A_2A_2
$A_1A_1 \times A_1A_1$	P^2	P^2	—	—
$A_1A_1 \times A_1A_2$	$2PH$	PH	PH	—
$A_1A_1 \times A_2A_2$	$2PQ$	—	$2PQ$	—
$A_1A_2 \times A_1A_2$	H^2	$\frac{1}{4}H^2$	$\frac{1}{2}H^2$	$\frac{1}{4}H^2$
$A_1A_2 \times A_2A_2$	$2HQ$	—	HQ	HQ
$A_2A_2 \times A_2A_2$	Q^2	—	—	Q^2
	总和	$\left(P+\frac{1}{2}H\right)^2$	$2\left(P+\frac{1}{2}H\right)\left(Q+\frac{1}{2}H\right)$	$\left(Q+\frac{1}{2}H\right)^2$
	=	p^2	$2pq$	q^2

复等位基因

当一个位点上有两个以上等位基因时，由基因频率确定基因型频率的方法，与两个等位基因的完全相同。如果 A_1 和 A_2 是任意两个等位基因，并且它们具有频率 q_1 和 q_2，那么处于哈迪—温伯格平衡之下的基因型频率如表 1-13 所示：

表 1-13 两个等位基因处于哈迪—温伯格平衡时的基因型频率

	基因型		
	A_1A_1	A_1A_2	A_2A_2
频率	q_1^2	$2q_1q_2$	q_2^2

这些频率也可通过一代随机交配而达到。

如果所有等位基因都是共显性的，这样所有基因型可在表型上识别，那么通过简单的计数便可从表型中估计基因频率。可是，如果一个或多个等位基因是隐性的，那么通过计数等位基因就不能获得基因频率。例如，在人类 ABO 血型系统中，O 等位基因对 A 和 B 是隐性的，因此 A 血型（表型）由两种基因型 AA 和 AO 组成，B 血型由 BB 和 BO 组成。

基因频率必须通过最大似然法来估计，正如 Weir（1990）所描述的事例那样。

性连锁基因

对于性连锁基因，其情形比常染色体基因要复杂得多。同配性别中基因频率和基因型频率之间的关系与一个常染色体基因相同，但异配性别只有两种基因型，且每一个体仅携带一个基因而不是两个。由于这个原因，群体中性连锁基因的三分之二由同配性别携带，三分之一由异配性别携带。为简明起见，异配性别将被指定为雄性。考察两个等位基因 A_1 和 A_2，频率为 p 和 q，假定基因型频率如表 1-14 所示：

表 1-14　性连锁的两个等位基因和基因型频率

	雌　性			雄　性	
	A_1A_1	A_1A_2	A_2A_2	A_1	A_2
频率	P	H	Q	R	S

于是雌性 A_1 的频率是 $p_f = P + \frac{1}{2}H$，雄性 A_1 的频率是 $p_m = R$。整个群体中 A_1 的频率是：

$$\left.\begin{aligned}\bar{p} &= \frac{2}{3}p_f + \frac{1}{3}p_m \\ &= \frac{1}{3}(2p_f + p_m) \\ &= \frac{1}{3}(2P + H + R)\end{aligned}\right\} \qquad [1.4]$$

现在，如果雄性和雌性之间的基因频率不同，则群体处于不平衡。整个群体中基因频率不会发生变化，但随着群体接近平衡，基因频率在两个性别之间的分布会来回摆动。其理由可从下面的考察中看出来。雄性只从它们的母亲那儿获得其性连锁基因；因此，p_m 等于上一代的 p_f。雌性从双亲那儿相等地获得其性连锁基因；因此，p_f 等于上一代的 p_m 和 p_f 的平均值。使用符号“′”表示后裔世代，有：

$$p_m' = p_f$$

$$p_f' = \frac{1}{2}(p_m + p_f)$$

两个性别中频率的差异为：

$$p_f' - p_m' = \frac{1}{2}(p_m + p_f) - p_f$$

$$= -\frac{1}{2}(p_f - p_m)$$

即为上一代差异的一半，但处于另一个方向。因此，两个性别之间基因的分布来回摆动，但在连续世代中差异减半，群体很快接近一个平衡，在平衡处两个性别的频率是相等的。当群体由一种类型的雌性（全为 A_1A_1）和另一种类型的雄性（全为 A_2）混合在一起而开始，并让它们随机繁殖时，图 1-2 说明了基因频率接近 2/3 时的平衡情况。

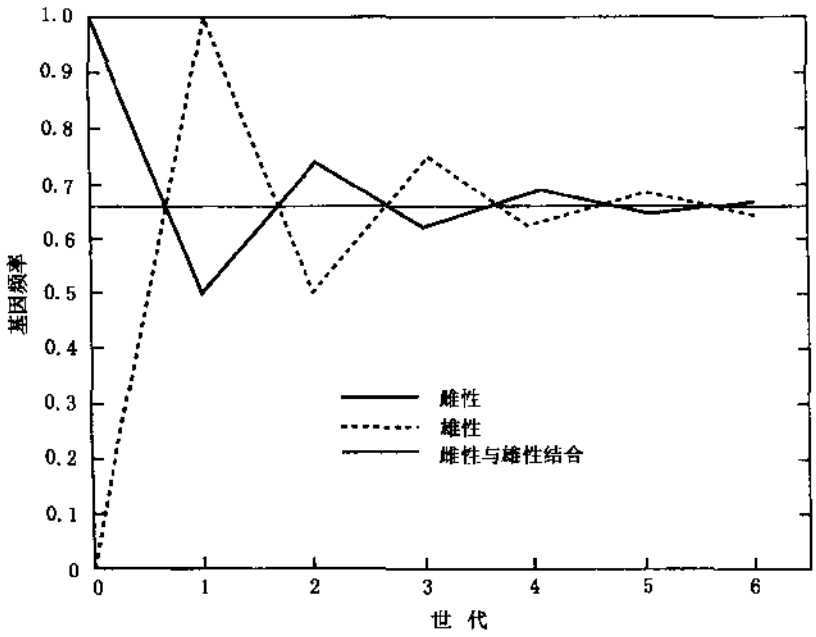


图 1-2　关于一性连锁基因在随机交配下接近平衡的情况

显示了雌性内、雄性内以及两性结合时的基因频率。这个群体从全是一种类型的雌性($q_f = 1$)和全是另一种类型的雄性($q_m = 0$)而开始的。

例 1.5

Searle (1949) 给出了伦敦猫一个样本中许多基因的频率。检查过的猫都是被送到诊所来杀死的；因此它们不一定是一个随机样本。在所研究的基因中间有以前称为“黄色”但现在叫“橙色”(O) 的性连锁基因。雌性中所有 3 种基因型都可识别，杂合体是“玳瑁壳”或“杂色”猫 (表 1-15)。数据用来对哈迪—温伯格期望值进行了测验，尤其要看是否有非随机交配的任何证据。第一个测验是要看两个性别中的基因频率是否相同。然后用与例 1.4 相同的方式对雌性基因型测验哈迪—温伯格法则。每一表型类别的数目以及从它们中计算出来的基因频率显示在下表中。雄性的基因频率稍高一点，但不显著。因此，到目前为止，没有理由认为群体不处于平衡。为简化起见，将对雌性观察到的基因频率视为用于计算雌性期望基因型频率的适宜基因频率。用与例 1.4 相同方式计算出的期望值给在下表中。观察到的数目不与期望值良好吻合，但与小期望值的不一致仅是值得怀疑的显著。如果不一致是真实的，则它可归因于非随机交配，但它也可能归因于人们对颜色的偏爱，这使得抽样有偏差，从而使样本不代表繁殖群体。对猫群体的更广泛的分析和讨论，见 Metcalfe 和 Turner (1971)。

表 1-15　应用伦敦猫的资料进行的哈迪—温伯格平衡 χ^2 测验

	个体数						
	雌　性				雄　性		
	++	+O	OO	总和	+	O	总和
观察数	277	54	7	338	311	42	353
期望值	273.4	61.2	3.4	338			
	$\chi^2_{[1]}=4.6$;	$P=0.04$					

	基因数			O 基因频率
	+	O	总和	q
雌性	608	68	676	0.101
雄性	311	42	353	0.119

多于一个位点

当分别考察所有常染色体位点时，经一代随机交配后基因型频率就达到平衡是真实的。但联合考察两个或更多个位点的基因型时，这不是真实的。为了说明这一点，假定有两个群体，一个完全由 $A_1A_1B_1B_1$ 基因型组成，另一个完全由 $A_2A_2B_2B_2$ 基因型组成。假定这两个群体以每一性别相等的数目进行混合，并允许随机交配。对两个位点的每一个位点上的两个等位基因来说，有 9 种可能的基因型，但只有 3 种将出现在第一代后裔中，它们是两个起始的双纯合体和双杂合体。由两个位点决定的性状之间存在完全的联合，两个性状似乎是由单个基因的差异所决定。随着连续的随机交配，丢失的基因型将在随后的世代中出现，但不会立即达到它们的平衡频率，并且性状之间起始的联合会逐渐减少。如果两个位点是连锁的，那么平衡频率的获得将花费更长时间，因为丢失的基因型出现取决于两个位点之间的重组。关于两个或更多位点的不平衡被称为配子期不平衡，或连锁不平衡，不管位点连锁与否。具有不同基因频率的群体混合或小群体内的偶然性可产生不平衡。选择较其他等位基因组合更为优势的等位基因组合也可产生并维持不平衡。表 1-16 是一个随机繁殖群体逼近平衡的速度的推导。

表 1-16　随机繁殖群体逼近哈迪—温伯格平衡速度的推导

基因	A_1	A_2	B_1	B_2
基因频率	p_A	q_A	p_B	q_B
配子类型	A_1B_1	A_1B_2	A_2B_1	A_2B_2
平衡时频率	p_Ap_B	p_Aq_B	q_Ap_B	q_Aq_B
实际频率	r	s	t	u
与平衡的差异	$+D$	$-D$	$-D$	$+D$

我们首先需要对不平衡数量的度量。这最好用配子类型的频率而不是合子基因型的频率来表示。考察两个位点，每一位点具有两个等位基因，基因频率显示在表 1-16 中。于是有 4 种类型配子。如果配子含有基因的随机组合，则群体处于平衡。因此，平衡时的配子频率只取决于基因频率，正如表 1-16 中显示的那样。假定实际的非平衡频率为 r、s、t 和 u，正如从表中看到的那样。这些中的每一个与平衡频率相差一个数量 D，两个配子

类型有正的离差，两个有负的离差。每一配子类型的 D 值必定是相同的，但符号除外。用 D 来度量不平衡的数量。通过比较相引和相斥双杂合体的频率、参照基因型可表示不平衡。不管两个位点连锁与否，基因型 A_1B_1/A_2B_2 可称为相引杂合体。它的频率是 $2ru$。相斥杂合体是 A_1B_2/A_2B_1 且它的频率是 $2st$。如果群体处于平衡，这两种基因型有相等的频率。与 D 的关系是：

$$D = ru - st$$

因此、D 等于相引和相斥杂合体间频率差异的一半。

例 1.6

黑腹果蝇（*Drosophila melanogaster*）的整个染色体可从自然群体的个体中抽取出来，并且可与含有经显性标记的平衡染色体的种蝇杂交，在这种染色体的抑制重组作用下，形成完全的纯合。一条染色体上任何位点数目的基因型被称为单倍型（haplotype），因为染色体基本上就是一种可能的单倍体配子类型。总数为 47 的第 2 号染色体从位于北卡罗来纳 Raleigh 的一个自然群体中抽取出来，用来评定 *scabrous* 基因区域内的限制性片段长度变异（Lai 等，1994）。下面的数据是两个 *Pst* Ⅰ限制位点每一个单倍型的染色体数目（表 1-17）。*Pst* Ⅰ(－12.0)（位点 *A*）定位在离 *scabrous* 基因转录起点的 5′ 端 12kb 处，*Pst* Ⅰ(5.8)(位点 *B*）则在离转录起点的 3′ 端 5.8kb 处。“＋”表示位点存在，“－”表示位点不存在。我们想要：(i) 确定这两个限制性位点是否处于连锁（不）平衡，(ii) 计算连锁不平衡系数 D。

表 1-17　两个 *Pst* Ⅰ限制位点每一单倍型的染色体数目

单倍型		观察数目	观察频率	期望频率	期望数目
A	*B*				
＋	＋	4	0.085	$p_Ap_B = 0.06$	2.8
＋	－	4	0.085	$p_Aq_B = 0.11$	5.2
－	＋	13	0.277	$q_Ap_B = 0.30$	14.1
－	－	26	0.553	$q_Aq_B = 0.53$	24.9

(i) 采用表 1-16 的术语，假定 p 表示每一个位点上＋等位基因的频率，q 代表－等位基因的频率。在两个限制性位点座位上基因频率的估计值是 $p_A = 8/47 = 0.17$；$q_A = 39/47 = 0.83$；$p_B = 17/47 = 0.36$；$q_B = 30/47 = 0.64$。于是，期望的单倍型频率和数目就像表 1-17 的右边显示的那样。

期望数目与观察数目相配得很好，差异在统计上是不显著的（$\chi^2_{[1]} = 0.93$；$P > 0.5$）。因此，对这些限制性位点，样本没有显示出群体处于连锁不平衡的证据。这些位点被 17.8kb 分开，这说明了非常紧密连锁的位点不一定处于连锁不平衡。

(ii) 上面给出了观察的单倍型频率，据此 $D = ru - st = (0.085) \times (0.553) - (0.085) \times (0.277) = 0.023$。$D$ 的绝对值取决于基因频率。D 的绝对范围从 -0.25 到 0.25；如果 $p_A = q_A = p_B = q_B = 0.5$，或 $r = u = 0$ 以及 $s = t = 0.5$（$D = -0.25$），或 $s = t = 0$ 和 $r = u = 0.5$（$D = 0.25$），这些值就出现。如果基因频率不是精确的相等，那么 D 的绝对值将在这个范围之

内。因此，D 经常用 D/D_{MAX} 表示，其中，D_{MAX} 是给定的观察基因频率下 D 所能取得的最大数值。表 1-16 中，$s = p_Aq_B - D \geq 0$ 和 $t = q_Ap_B - D \geq 0$，所以，D 一定 $\leq p_Aq_B$ 且 $\leq q_Ap_B$。因此，D_{MAX} 是 p_Aq_B 或 q_Ap_B 中较小的一个。对这个例子，D_{MAX} 是 0.11，因此 $D/D_{MAX} = 0.21$。D 的计算值是它最大值的 21%，它是重要的。大的样本大小对检测连锁不平衡是必要的，特别是当基因频率处于极端值时。

当处于连锁不平衡的群体随机交配时，不平衡的数量会随着每一个后续的世代而逐渐降低。这种情况发生的速率取决于两个连续世代中配子类型的频率。如果两个位点被认为在同一条染色体上连锁，那么这种情况可能最容易具体化。后裔世代中不平衡系数 D，可依 4 种配子类型的任一种的频率而获得，因此，让我们只考察 A_1B_1 类型。这种类型可以两种方式出现于后裔配子中。首先，它可作为基因型 A_1B_1/A_xB_x 的非重组体而产生，下标 x 意味着两个等位基因中任一个都可出现。以这种方式产生的 A_1B_1 的频率是 $r(1-c)$，r 是亲代配子中 A_1B_1 的频率，c 是重组频率。其次，或者它可作为基因型 A_1B_x/A_xB_1 的重组体而产生。A_1B_x 染色体的频率是 p_A，A_xB_1 染色体的频率是 p_B。那么，以这种方式产生 A_1B_1 的频率是 p_Ap_Bc。因此，后裔配子中 A_1B_1 的频率是：

$$r' = r(1-c) + p_Ap_Bc$$

并且后裔世代中的不平衡系数是：

$$\begin{aligned} D' &= r' - p_Ap_B \\ &= r(1-c) - p_Ap_B(1-c) \\ &= (r - p_Ap_B)(1-c) \\ &= D(1-c) \end{aligned}$$

如果我们再考察一个世代的这种过程，我们得到：

$$D'' = D'(1-c) = D(1-c)^2$$

这样，在任何数目 t 世代之后，不平衡系数由下式给出：

$$D_t = D_0(1-c)^t \qquad [1.5]$$

不平衡中的位点不必是连锁的。对不连锁位点 $c = \frac{1}{2}$，不平衡的数量随着每一世代的随机交配而减半。对连锁位点，不平衡消失更缓慢。图 1-3 显示了不同程度的连锁在 12 个世代过程中不平衡数量是如何减少的。

只要（1 − c）定义为任何位点之间没有重组时一个配子传递一代的概率，由上面等式给出的平衡途径就同等地适用于联合考察的任何数目位点的不平衡。位点数目越大，没有重组的概率就越小；对两个不连锁位点，它是 1/2，对 3 个不连锁位点，它是 1/4，对 4 个不连锁位点，它是 1/8。这样，多位点不平衡比 2 个位点衰退得更快，不久就控制了许多位点间的总的不平衡。这种情形的实际后果是当许多位点可用于研究时，不平衡更象是在成对位点中发现的，而不是在联合考察的较多数目的位点中发现的。有关 3 个位点的细节见 Crow 和 Kimura（1970），其估计方法见 Weir 和 Cockerham（1979）。

在自然群体中经常发现重组来不及消除起始不平衡位点之间存在的连锁不平衡，例如疾病位点和紧密连锁的 RFLP 之间。找到一个多态位点和一个疾病基因之间完全联合（完全连锁不平衡）实际上是在分子水平上鉴定疾病基因特性的第一步，并且对遗传诊断非常有用。

可是，紧密连锁并不自动意味着位点将是处于连锁不平衡之中。甚至紧密连锁位点却可能处于连锁平衡之中（见例 1.6）；如果位点上具有不同基因频率的两个群体混合在一起，完全不连锁的位点能够处于连锁不平衡，正如上面所例举的那样。因此，对群体不平衡的了解能够告诉我们有关群体繁育史的一些情况，单个位点不能做到这一点，因为它们经单个世代随机交配之后便达到哈迪—温伯格频率（Crow，1986）。连锁不平衡的其他原因是小群体中不同位点等位基因之间的偶然联合，以及对特定等位基因组合的选择。连锁不平衡具有的重要性，在以后各章尤其是关于数量性状的选择中必须加以考虑。

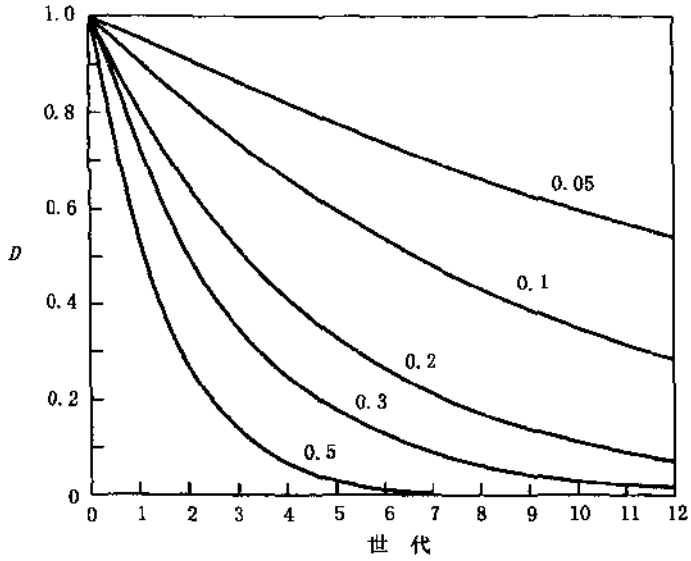


图 1-3 联合考察的两个位点随机交配之下接近平衡的情况

曲线显示了相对于 0 世代不平衡的不平衡数量 D。5 条曲线指的是两个位点之间不同的连锁程度，它们出沿着每条曲线显示的重组频率来表示。标明 0.5 的曲线指的是不连锁位点。

非随机交配

有两种不同形式的非随机交配。第一种是交配个体因祖先血统而彼此有亲缘。这往往增加所有位点上的纯合体频率。它的效应将在第 3 章中描述。第二种是个体针对所考察的任一特定位点上的基因型倾向于优先进行交配。在此需简要地论及这种形式的非随机交配。

选型交配

如果配偶具有相同表型的机会比偶然发生的更频繁，就被称为选型交配，如果不频繁则被称为非选型交配。在表型反映基因型的程度上，选型或非选型交配影响基因型频率。Crow 和 Kimura（1970）描述了这种影响，这里只简要列出。选型交配在人类群体中具有一些重要性，选型交配针对身高、智力和其他性状而发生。可是，这些不是单一的基因差异，不属本书讨论的范围。非选型交配在植物的自体受精体系中是普遍的。

通过适当修改交配类型的频率以允许增加相似表型之间的交配频率，可从表1-12中推

导出关于单个位点选型交配的后果。对后裔之间基因型频率的影响是增加纯合体的频率和减少杂合体的频率。实际上，群体只要部分地细分为两组，交配在组内就比在组间更经常发生。如果选型交配在连续世代中继续下去，那么群体接近一种平衡状态，在该状态下基因型频率保持不变。

非选型交配的后果，一般地说与选型交配的相反：它引起杂合体的增加和纯合体的减少。可是，非选型交配通常具有改变基因频率的额外后果。如果交配主要在不同表型之间进行，那么与普通表型相比，较稀少的表型在交配时有更多的成功机会。因此，较稀少的等位基因占优势，基因频率趋向中间值，在这个数值上，表型在频率上是相等的。一个熟悉的非选型交配例子是两性生殖方式，它直接导致频率为0.5的一个基因或染色体。植物的自体受精机制以复等位基因为基础，稀有等位基因的有利地位导致了大量的等位基因均以大致相等的频率并存。

习 题

1.1 在美国白人的一个样本中记录了人类M-N血型的下列数字。

M	MN	N
1 787	3 039	1 303

(1) 在这个样本中观察到的基因型频率是多少?

(2) 基因频率是多少?

(3) 对观察到的基因频率，按哈迪—温伯格法则预测的基因型频率是多少?

(4) 观察到的频率和期望值一致性程度怎样?

数据来自 Wiener, A.S. (1943) 由 Stern, C. (1973) 引用, *Principles of Human Genetics*. Freeman, San Francisco。

[解答序号: 1]

1.2 大约30%的人不能辨别苯氨基硫甲酸 (PTC) 的苦味。没有能力辨味是由于常染色体上一单个隐性基因引起的。假定群体处于哈迪—温伯格平衡，那么不能辨味基因的频率是多少?

[解答序号: 11]

1.3 白化病在欧洲人群体中发生的频率大约是每20 000人中有1人。假定它是由常染色体上一单个隐性基因引起的，并假定群体处于哈迪—温伯格平衡，那么携带者的比例是多少? 只需给出一个近似答案。

[解答序号: 21]

1.4 作为一种代数练习，计算出将导致$\frac{1}{3}$的正常个体为携带者的一个随机交配群体中一个隐性突变体的基因频率。

[解答序号: 31]

1.5 在178名英国人的一个样本中发现了红细胞酸性磷酸酯酶的3个等位变异A、B和C。通过电泳可识别所有基因型，并且样本中的频率是:

基因型	AA	AB	BB	AC	BC	CC
频率 (%)	9.6	48.3	34.3	2.8	5.0	0.0

样本中基因频率是多少？为什么没有发现 CC 个体？

数据来自 Spencer，N.，等（1964）*Nature*，201，299～300。

［解答序号：41］

1.6 由于一个性连锁隐性基因的缘故，大约 7% 的男人是色盲。假定处于哈迪—温伯格平衡，(1) 女人是携带者的预期比例是多大？(2) 女人是色盲的预期比例是多大？(3) 以什么比例结婚，预期丈夫和妻子都是色盲？

［解答序号：51］

1.7 无眼（*so*）和朱砂（*cn*）是黑腹果蝇（*Drosophila melanogaster*）常染色体上的两个隐性基因。它们紧密连锁且能把它们作为一个位点上的等位基因来看待。“杂合体” *so*/*cn* 是野生型，可与两种纯合体区分开来；(如果另一种眼睛颜色突变体即褐色（*bw*）突变体种蝇被纯合时，则 *so*/*so* 没有眼睛；*cn*/*cn* 有白眼睛)。在一个分类实验中，来自种蝇 *so*/*so* 的 4 个雄性和 4 个雌性与来自种蝇 *cn*/*cn* 的 16 个雄性和 16 个雌性一起放入一个小瓶并让它们交配。有 20 个这样的小瓶。通过基因型分类其后裔总数如下：

so/*so*	*so*/*cn*	*cn*/*cn*
135	359	947

这些数字与哈迪—温伯格期望值有怎样的不同？提出不一致的理由。

［解答序号：61］

1.8 假定果蝇（*Drosophila*）被安置在习题 1.7 所描述的小瓶中繁殖，但这次基因频率为 0.5。通过在每个小瓶中放入每一种蝇的 10 个雄性和 10 个雌性可做到这一点。*so*/*so* 雌性的供应用完了，只剩下 4 个用于最后一个小瓶。因此，为了保持计划的基因频率和亲本数，这个小瓶组成如下：16♂♂ + 4♀♀的 *so*/*so* 和 4♂♂ + 16♀♀的 *cn*/*cn*。得到这个小瓶的学生对他发现的情形有点惊奇。你期望的后裔中基因型频率是多少？

［解答序号：71］

1.9 证明当一个位点上存在任何数目的等位基因，并且所有等位基因具有相同的频率时，杂合体的总频率最大。那时杂合体的总频率是多大？

［解答序号：81］

1.10 假定基因型 AABB 的一个品系和基因型 aabb 的另一个品系混合在一起，两个品系的数目相等，且雄性和雌性的数目也相等，它们随机交配。将这种亲本世代称为 0 世代。随后的世代还是随机交配并且基因型之间没有生殖力和生活力的差异。如果两个位点 (1) 不连锁，(2) 具有 20% 重组频率的连锁，那么 2 世代即重组 2 代后的后裔中基因型 AAbb 的频率将是多少？

［解答序号：91］

1.11 如果取两品系中一个品系的雄性和另一个品系的雌性进行杂交，那么习题 1.10 的解答会怎样改变？

［解答序号：101］

第 2 章　基因频率的变化

我们已经看到一个随机交配的大群体，如果不存在改变其遗传特性的作用，则它的基因频率和基因型频率是稳定的。现在我们可以继续研究这些作用，通过它们使基因频率，进而也使基因型频率发生变化。有两类过程：*系统过程*，它倾向于以一种在大小和方向上都可预测的方式改变基因频率；*离散过程*，它是在小群体中由抽样效应产生的，只能在大小上但不能在方向上加以预测。本章中我们只关注系统过程，而且为了将离散过程从模型中排除出去，我们将只考察随机交配的大群体。有 3 种系统过程：迁移、突变和选择。开始我们将分别研究这些过程，假定一次只有一种过程起作用，然后我们将看看不同的过程如何相互作用。

迁　移

迁移的效应在这里可很简单地加以处理，仅略略提一下，虽然以后在谈到小群体时，我们将更多地讨论它。让我们假定在每一世代中都由比例为 m 的新迁入者组成一个大群体，剩余的 $1-m$ 为本地品种。假定迁入者中某一基因的频率为 q_m，本地品种中该基因的频率为 q_0。于是混合群体中该基因的频率 q_1 将是：

$$\begin{aligned} q_1 &= mq_m + (1-m)q_0 \\ &= m(q_m - q_0) + q_0 \end{aligned} \qquad [2.1]$$

由一个世代的迁移引起的基因频率变化 Δq，是迁移前的频率和迁移后的频率之差。因此

$$\begin{aligned} \Delta q &= q_1 - q_0 \\ &= m(q_m - q_0) \end{aligned} \qquad [2.2]$$

这样，受迁移影响的群体中基因频率的变化率很显然取决于迁移率以及迁入者和本地品种之间基因频率的差异。

突　变

根据我们所关注的突变是非常稀有以至于实际上是独一无二的一个事件，还是会频频发生的一个步骤，突变对群体遗传特性的影响有所不同。前者在一个大群体中不产生永久性的变化，而后者则产生永久性的变化。

非频发突变

首先考虑一个突变事件，它在整个群体中仅产生单个突变基因或染色体的一种代表。这类突变作为基因频率变化的一种原因，一点也不重要，因为在一个大群体中一罕见突变的产物仅有很小的存活机会。最初出现在杂合体中的突变基因，于下一世代内它被丢失的机会是 1/2。如果它生存下来，它可能被一个或多个拷贝所体现，但每一个拷贝只有一半的机会可存活到第三个世代。丢失是永久性的，因此无限期生存的机会实际上很小，在一

个无限大的群体中为0。因为实际群体不可能无限大，一定得把罕见突变预料成非常偶然地无限期生存下去，并导致基因频率的变化。关于这一点在本章的后面以及第4章中将更多地谈到。

频发突变

正是第二种类型的突变——频发突变，才是我们主要关注的作为造成基因频率变化的一种作用，并且在一个大群体中，一个突变基因的频率决不会这么低以致于由于抽样而发生完全丢失。于是，我们必须查明突变的这种“压力”对群体基因频率将产生什么样的影响。

假定基因 A_1 以每世代的频率 u（u 是一个世代和下一世代间突变成 A_2 的所有 A_1 基因的比例）突变成 A_2。如果在一个世代中 A_1 的频率是 p_0，那么在下一个世代中新突变生成的 A_2 基因的频率是 up_0。因此，A_1 的新的基因频率是 $p_0 - up_0$，基因频率的变化是 $-up_0$。现在考察当基因在两个方向上突变时会产生什么结果。为简单起见，假定只有两个等位基因 A_1 和 A_2，具有起始频率 p_0 和 q_0。A_1 以每世代速率 u 突变成 A_2，A_2 则以每世代速率 v 突变成 A_1。于是一个世代之后，A_2 基因有一个增量，在一个方向上发生突变时，它等于 up_0，在另一个方向上发生突变时，便有一个损失，它等于 vq_0。用符号表示，有这样的情况：

$$\begin{array}{ll} \text{突变率} & A_1 \underset{v}{\overset{u}{\rightleftharpoons}} A_2 \\ \text{起始基因频率} & p_0 \quad\quad q_0 \end{array}$$

那么一个世代中基因频率的变化是：

$$\Delta q = up_0 - vq_0 \qquad [2.3]$$

容易看出这种情况导致基因频率处于一种平衡状态，在该频率时不会发生进一步的变化，因为如果一个等位基因的频率有了增加，同一方向突变的剩余等位基因频率就会低些，而反方向的频率就会高些。令频率变化 Δq 等于0可以找到平衡点。因此在平衡时有：

$$\left.\begin{array}{ll} & pu = qv \\ \text{或} & \dfrac{p}{q} = \dfrac{v}{u} \\ \text{和} & q = \dfrac{u}{u+v} \end{array}\right\} \qquad [2.4]$$

从突变对基因频率的影响中可得出两个结论。突变率一般很低，对大多数生物体的大多数位点来说，每世代大约为 10^{-5} 或 10^{-6}。这意味着在任一特定位点上约有十万分之一至百万分之一的配子携带一个新的突变生成的等位基因。因此，对正常的突变率，单靠突变只能产生很慢的基因频率变化；在进化的时间尺度上，它们可能是重要的，但除微生物之外，它们几乎不能通过实验检测出来。第二个结论涉及两个方向上突变之间的平衡。对回复突变（从突变型到野生型）的研究表明，它经常比向前突变（从野生型到突变型）要少得多（Muller 和 Oster，1957；Schlager 和 Dickie，1971）。如果回复突变只有向前突变的 1/10，那么单独由突变引起的平衡基因频率将是：野生型等位基因为0.1，突变型等位基因为0.9；换句话说，突变型将是普遍的类型，而“野生型”是罕见的类型。因为这不是在自然群体中发现的情形，那么很明显，这类基因的频率不单单是突变的产物。在下一节中

我们将看到突变等位基因的稀少性可归因于选择。

选 择

到目前为止，我们假定群体中所有个体对下一代都有相等的贡献。现在我们必须考虑个体在存活力和生殖力方面有所不同的事实，因此它们给下一个世代贡献不同数目的后代。各个体对下一世代的后代贡献被称为个体的适合度（fitness），或有时称为适应值（adaptive value）或选择值（selective value）。如果适合度的差异以任何方式与个体基因型中一特定基因的存在或缺失相关联，那么选择对那个基因可以起作用。当一个基因受到选择时，它在后代中的频率与在亲本中的不同，因为不同基因型的亲本不等量地传递它们的基因给下一世代。通过这种方式，选择引起了基因频率的变化，进而也导致了基因型频率的变化。由选择引起的基因频率的变化描述起来比由突变引起的更复杂，因为产生于选择作用的适合度差异是表型的一个方面。因此我们必须考虑由所讨论的基因表现出的显性度。在这方面，显性意味着适合度的显性，这不一定与基因的主要可见效应的显性相同。例如，大多数突变基因在其可见效应中对野生型是完全隐性的，但这不一定意味着杂合体具有与野生型纯合体相等的适合度。我们将要涉及的不同显性度的含意于图2-1给予说明。

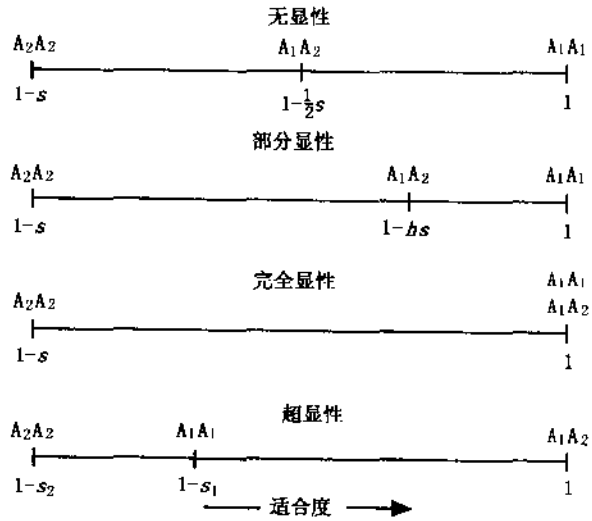


图 2-1 关于适合度的显性度

只要一种或另一种基因型携带我们所讨论的基因，那么采用选择淘汰基因型的方式来考察对该基因不利的选择最方便。这可以通过降低存活力或通过降低广义生殖力（包括交配能力），或通过两者来操作。在个体的生活史中，选择首先通过存活力，然后通过生殖力起作用。因此，我们必须推导出从一个世代的合子到后裔世代的合子的基因频率变化。基因频率在合子中不能观察到，因此从观察到的基因频率变化中推导出选择力量存在实际困难，但我们在后面将回到这些上来。选择的强度可用选择系数 s 表示，它是一种特定基因型与一种标准基因型（通常是最有利的）相比较配子贡献减少的比率。有利基因型的贡

献取 1，于是不利基因型的贡献就是 $1-s$。这表示了一种基因型相对于另一种基因型的适合度。例如，假定选择系数是 $s=0.1$；那么适合度是0.9，这意味着有利基因型的每产生100个合子，对于不利的基因型只产生90个。以对子代贡献率定义的适合度，严格地说应称为相对适合度，但在以后的论述中，都将它称为适合度。

涉及任何特定位点的基因型适合度，在个体间不必相同。它取决于个体生活的环境条件，还取决于其他位点上基因的基因型。当我们给某一基因型指定一确定的适合度时，这指的是这种基因型在整个群体中的平均适合度。虽然个体间适合度的差异导致选择可同时作用于很多位点，甚至全部位点，但这里我们将把注意力限制在选择对单个位点上基因的影响上，并假定不同基因型的平均适合度保持稳定不变，而不管由同时作用于其他位点的选择所引起的变化。我们所获得的结论适用于没有人类干涉的自然条件下发生的自然选择，也同样适用于人工选择，即育种者或实验者对亲本个体的选择以及他对从每一个亲本中选择饲养的后代的数目所做的选择。

选择下基因频率的变化

首先我们必须推导出由一个世代的选择引起的基因频率变化的基本公式。然后才可以考虑关于选择的有效性，这些公式能告诉我们些什么。必须考察各种显性的情况，但方法都是相同的，将借助完全显性的一种情形来说明它，其中的选择作用不利于隐性纯合体。表 2-1 显示了选择之前基因型具有的哈迪—温伯格频率。A_2A_2 是隐性纯合体，具有对它起不利作用的选择系数 s。下一行给出了每一种基因型的适合度。将起始频率乘以适合度就给出了选择后每一种基因型的频率。这作为“配子贡献”被列入，以便考察选择在整个生活史中起的作用。注意选择后总频率已不再是 1，因为由于选择，已有等于 sq^2 的损失部分。为了找出所产生的 A_2 配子的频率，即后代中 A_2 基因的频率，我们取 A_2A_2 个体的配子贡献，加上 A_1A_2 个体的配子贡献的一半，再除以新的总频率，即我们应用等式 [1.1]。因此新的基因频率是：

$$q_1=\frac{q^2(1-s)+pq}{1-sq^2}$$

通过代入 $p=(1-q)$ 可以简化这个式子。然后重新整理得出：

$$q_1=\frac{q-sq^2}{1-sq^2} \qquad [2.5]$$

表 2-1 对一隐性基因的不利选择

	基因型			
	A_1A_1	A_1A_2	A_2A_2	总和
起始频率	p^2	$2pq$	q^2	1
选择系数	0	0	s	
适合度	1	1	$1-s$	
配子贡献	p^2	$2pq$	$q^2(1-s)$	$1-sq^2$

由一个世代的选择引起的基因频率的变化 Δq 为：

$$\Delta q=q_1-q$$

代入等式［2.5］中的 q_1，经过一些重新整理之后，变成：

$$\Delta q = -\frac{sq^2(1-q)}{1-sq^2} \qquad [2.6]$$

从这个式子我们看出，选择对基因频率的影响不仅取决于选择强度 s，而且取决于起始基因频率。但这两者关系都有些复杂，对其显著性检验直到其他情况已被讨论后才能进行。

对不同显性情况的新的基因频率和基因频率变化的表达式给在表 2-2 中。表中一般表达式（5）允许对具相应适合度的任何显性度 Δq 进行求解。两个表达式（3 和 4）是根据选择的方向，为一个完全显性的基因提供的。第一个，它是上面推导出来的，适用于不利于隐性纯合体的选择。相反，如果选择不利于显性的表型，Δq 就会很不相同。通过考察完全淘汰的后果，此时 $s=1$，可以最充分地注意到这种差异。于是对显性表型不利选择的公式就简化成 $q_1=1$，它表示了这一事实：如果只有隐性纯合体繁殖，基因频率立即就达到 1。另一方面，隐性纯合体的完全淘汰将留下存在于杂合体中的所有隐性基因。随着 s 值的降低，相反方向选择效应之间的差异变得较不显著。到目前为止，提到的所有形式的选择最终倾向于从群体中淘汰某一个等位基因。相反，适合度的超显性，在此杂合体的适合度占优势，倾向于在群体中保持两个等位基因。对这种形式的选择以后要较详细地加以讨论。

表 2-2　在适合度显性不同的情况下，由一个世代的选择所引起的基因频率变化（A_2 的起始基因频率是 q）

	基因型的起始频率和适合度			新的基因频率	基因频率的变化
	A_1A_1 p^2	A_1A_2 $2pq$	A_2A_2 q^2	q_1	$\Delta q = q_1 - q$
(1)	1	$1-\frac{1}{2}s$	$1-s$	$\frac{q-\frac{1}{2}sq-\frac{1}{2}sq^2}{1-sq}$	$-\frac{\frac{1}{2}sq(1-q)}{1-sq}$
(2)	1	$1-hs$	$1-s$	$\frac{q-hspq-sq^2}{1-2hspq-sq^2}$	$-\frac{spq[q+h(p-q)]}{1-2hspq-sq^2}$
(3)	1	1	$1-s$	$\frac{q-sq^2}{1-sq^2}$	$-\frac{sq^2(1-q)}{1-sq^2}$
(4)	$1-s$	$1-s$	1	$\frac{q-sq+sq^2}{1-s(1-q^2)}$	$+\frac{sq^2(1-q)}{1-s(1-q^2)}$
(5)	$1-s_1$	1	$1-s_2$	$\frac{q-s_2q^2}{1-s_1q^2-s_2q^2}$	$+\frac{pq(s_1p-s_2q)}{1-s_1p^2-s_2q^2}$

(1) 无显性；选择不利于 A_2。

(2) A_1 部分显性；选择不利于 A_2。

(3) A_1 完全显性；选择不利于 A_2。

(4) A_1 完全显性；选择不利于 A_1。

(5) 超显性；选择不利于 A_1A_1 和 A_2A_2（也适用于相对于 A_1A_2 表达的适合度的任何显性度）。

表 2-2 中 Δq 的表达式相当繁琐，借助足以对多种目的都可靠的一种近似方法来简化它们，经常是有用的。如果选择系数 s 或者基因频率 q 很小，那么表 2-2 中等式的分母变得非常接近 1，我们可以单独使用分子作为 Δq 的表达式。于是，对任一方向的选择，无显性时，有：

$$\Delta q = \pm \frac{1}{2} sq(1-q) \quad （近似） \qquad [2.7]$$

具有完全显性时：

$$\Delta q = \pm sq^2(1-q) \quad （近似） \qquad [2.8]$$

选择的有效性

从这些公式中可看出选择的有效性即 Δq 的大小取决于起始基因频率 q。这种关系的性质可从显示不同 q 值的 Δq 的曲线中得到最充分地了解。图 2-2 显示了无显性和完全显性 情形下的这些曲线。它们在两个方向的选择之间也有所区别。选择一个值 $s = 0.2$ 作为

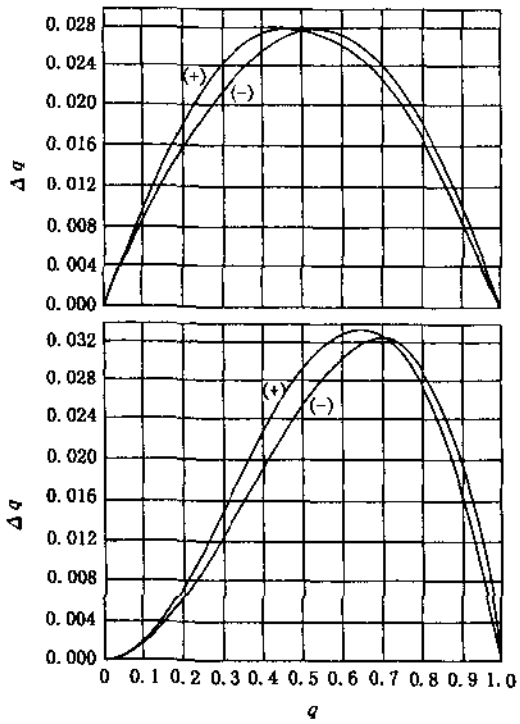


图 2-2　在不同起始基因频率 q 值和选择强度 s = 0.2 的情况下基因频率的变化 Δq

上图：一个无显性的基因。下图：一个完全显性的基因。标有（-）号的曲线指的是不利于频率为 q 的基因的选择，因此 Δq 是负的。标有（+）号的曲线指的是有利于这个基因的选择，因此 Δq 是正的（根据 Falconer，1954）。

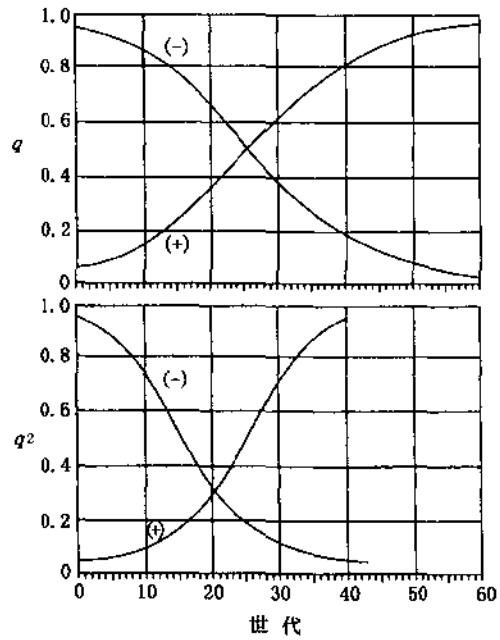


图 2-3　选择过程从一个极端到另一个极端的期间基因频率的变化（选择强度 s = 0.2）

上图：一个无显性的基因。下图：一个完全显性的基因，q 是隐性等位基因的频率，q^2 是隐性纯合体的频率。标有（-）号的曲线指的是不利于频率为 q 的基因的选择，因此 q 或 q^2 下降。标有（+）号的曲线指的是有利于该基因的选择，因此 q 或 q^2 增加（根据 Falconer，1954）。

选择系数，因为这似乎是在实验室选择实验中对与数量性状有关的基因起作用的选择系数的正确数量级，其理由在第 12 章给出。首先，我们可以注意到，对这个 s 值，根据选择方向，Δq 决不会有大的差异。这些曲线证明的关于选择有效性的两个要点是：(1) 在中间基因频率时选择最有效，当 q 很大或很小时，选择最无效；(2) 当隐性等位基因稀少时，有利于或不利于一个隐性基因的选择都是极其无效的。这是前面提到的这个事实的后果：当一个基因稀少时，它几乎完全存在于杂合体中。

观察起始基因频率对选择有效性影响的另一种方法，是从一个或另一个极端开始作条曲线，表示多世代中的选择过程。这样的曲线显示在图 2-3 中。它们直接从图 2-2 来构建，选择系数仍取 $s=0.2$。它们表明由选择引起的变化开始时很慢，无论它是从高的还是低的起始基因频率开始；在中间频率时它变得很快，趋于频率极端时又变慢。在完全显性基因的情形中，人们主要对纯合隐性基因型的频率即 q^2 感兴趣。由于这个原因，这条曲线显示了选择对 q^2 而不是对 q 的影响。

例 2.1

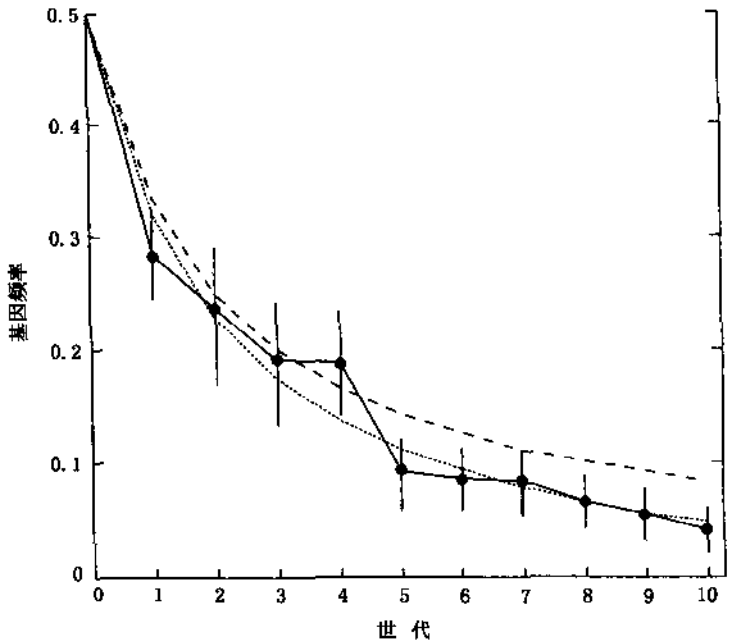


图 2-4　根据例 2.1 的描述，实验室中自然选择下的基因频率变化

(引自 Wallace，1963)

图 2-4 显示了 Wallace (1963) 描述的黑腹果蝇 (*Drosophila melanogaster*) 中一常染色体隐性致死基因的基因频率变化。这个群体从全部是杂合体的果蝇开始，因而 0 世代的基因频率是 $q=0.5$。随后每一世代的亲本都是前一世代存活后裔的一个随机样本。只有杂合体和正常纯合体能存活。杂合体通过测交来识别。每一世代中测验大约 100 到 200 个果

蝇，这给出了总数大约为200到400个基因，从中来估计基因频率。观察到的基因频率±2倍标准误对10个连续世代的每一代进行绘图。

通过表2-2中 q_1 的公式，计算了每一世代的期望基因频率。计算了两个期望值。第一个，由虚线所表示，假设致死基因完全隐性。对 $s=1$，表2-2中（3）行的公式简化成 $q_1=q/(1+q)$。观察结果表明对一个完全隐性致死基因来说，基因频率比预期的要减少得稍快一点。由点线表示的第二个期望值假设杂合体的适合度减少10%。对 $s=1$，表2-2中（2）行的公式简化成 $q_1=pq(1-h)/[p^2+2pq(1-h)]$，对 $h=0.1$，计算了这个算式的值。结果与这个期望值正好符合。

需要的世代数

需要多少世代的选择才可以影响基因频率使它产生一个特定的变化？这个问题的回答可能涉及育种方案或提出的优生措施。我们只考虑选择不利于隐性基因，即不希望的纯合体完全被淘汰，$s=1$ 的情况。这也适用于不利于一个隐性致死基因的自然选择，以及在一个育种方案中不利于一个不想要的隐性基因的人工选择。暂时我们还要假定没有突变。一个世代之后新的基因频率的表达式在等式［2.5］中（以及表2-2的（3）行中）给出。

将 $s=1$ 代入这个等式中，将0，1，2，…，t 世代选择之后的基因频率记为 q_0，q_1，q_2，…，q_t，有：

$$q_1=\frac{q_0}{1+q_0}$$

和

$$q_2=\frac{q_1}{1+q_1}$$

$$=\frac{q_0}{1+2q_0}$$

通过代入 q_1 和简化而得到上式。因此一般地：

$$q_t=\frac{q_0}{1+tq_0} \qquad [2.9]$$

将基因频率从 q_0 改变到 q_t 所需要的世代数 t 是：

$$t=\frac{q_0-q_t}{q_0q_t}$$

$$=\frac{1}{q_t}-\frac{1}{q_0} \qquad [2.10]$$

下面的例子说明了已经提出的这一点，即当一个隐性基因的频率低时，要靠选择作用去改变它是很缓慢的。

例2.2

有时有人建议作为一种优生措施，患有严重遗传缺陷的那些个体应该避免生育，因为通过这种方式，这类缺陷的频率将在未来的世代中就会降低。在确定这是否是一条好建议之前，我们应该知道这样做可以期望获得什么结果。不考虑突变我们就不能恰当地讨论这个问题，如同在以后要谈到的一样；我们忽略突变得到的答案，用来表现我们希望的最佳

情况。白化症将用作一个事例，虽然它不是一个很严重的缺陷。假定白化症由一单个隐性基因引起，那么将白化个体的频率降低到目前值的一半将花费多长时间？欧洲人群内目前的频率大约是1/20 000。这是 q_0^2，它给出 $q_0 = 1/141$。目标是 $q_t^2 = 1/40\ 000$，$q_t = 1/200$。这样，根据等式 [2.10]，$t = 200 - 141 = 59$ 个世代。一个世代 25 年，它将花费近1 500年才能这到这个适度的目标。实际上，白化症遗传实体不是一单个基因，而是至少由两个不同的隐性基因所引起，每一个的频率都低于 1/141。因此实际上淘汰将更缓慢。

当然在家养物种中，有害基因的淘汰可通过后裔测验大大加速。与已知的杂合体进行测交，这可鉴别出所测验个体之间的杂合体。于是基因可很快被淘汰。基因仅存留在因后裔数目不充足而被误归类为正常个体的杂合体中。

平均适合度和负荷

当基因频率被选择改变时，一些个体由于它们不能生存或繁殖而一定遭受了“遗传死亡”，因此群体的平均适合度降低了。遭受遗传死亡的群体比例被称为负荷（load），它是由于群体中存在有害基因的缘故而为群体所携带。如果 L 是负荷，那么群体的平均适合度是 $1 - L$。平均适合度和负荷在表 2-1 中进行了推导，但没有特别指明。平均适合度是选择后基因型频率的总和，是表 2-2 中 q_1 或 Δq 所有表达式的分母。例如，对一个隐性基因，平均适合度是 $1 - sq^2$，那么负荷是 sq^2。平均适合度仍然是相对适合度，相对于其中没有有害基因的一个群体。负荷对群体不一定是一个实际的有害物，因为大多数物种产生的后裔比其环境资源能承受的要多，由遗传原因引起的一些个体的死亡可为其他个体腾出空间，否则它们会因缺乏食物或一些其他原因而死亡。例如，有一种果蝇（*D. tropicalis*，来自中美洲），在某一地区有 50% 的个体遭受了遗传死亡，但群体仍然繁殖得很旺盛（Dobzhansky 和 Pavlovsky，1955）。

平　衡

突变和选择之间的平衡

我们已分别描述了突变和选择的效应，现在必须比较它们并联合起来考察它们。在引起基因频率的变化中，哪一个是更有效的过程？将我们在自然群体中发现的有害基因低频率，归因于倾向增加频率的突变和倾向减少频率的选择之间的平衡，这合理吗？已经获得的在突变或选择单独作用下的基因频率变化的表达式，表明两者都取决于起始基因频率，但方式不同。当突变基因稀少时，对一特定基因来说，增加其频率，突变是最有效的（因为有更多未突变的基因要突变）；但是当基因稀少时，选择最无效。因此两种过程的相对有效性取决于基因频率，并且如果两个过程起作用的时间足够长，则最终将达到一种平衡状态。因此，我们必须找出达到平衡时，基因频率将是多大。通过使基因频率变化的两个表达式相等可以做到这一点，因为平衡时，由突变引起的变化将与由选择引起的变化大小相等而方向相反。

让我们考察频率为 q 的一个基因（A_2），向它的突变率为 u，离它的突变率为 v，不利于它的选择系数是 s。于是根据等式 [2.3] 和表 2-2 中具有部分显性的基因频率变化的

表达式，在平衡处有：

$$up - vq = \frac{spq[q + h(p - q)]}{1 - 2hspq - sq^2} \qquad [2.11]$$

这个等式太复杂以致于对我们的问题不能给出一个清晰的答案。但我们只须在准确度上略微牺牲就能够作出两点简化。我们对在低平衡频率处的基因特别感兴趣。如果 q 很小，代表回复突变的一项 vq 相对不重要，可以忽略不计；我们可以使用近似表达式表示选择效应（即设定分母为1）。于是近似地，$up = spq\,[q + h\,(p - q)]$。对于 A_2 隐性（$h = 0$）、A_2 无显性（$h = 0.5$）和 A_2 显性（$h = 1$）的特殊情形，这个表达式可以进一步简化。首先考察不利于一个完全隐性基因的选择的平衡条件：

$$u = sq^2 \qquad \text{（近似）} \qquad [2.12]$$

$$q = \sqrt{\frac{u}{s}} \qquad \text{（近似）} \qquad [2.13]$$

对无显性的基因，表达式简化为 $u = 0.5sq$（近似），因此平衡时有：

$$q = \frac{2u}{s} \qquad \text{（近似）} \qquad [2.14]$$

对于不利于一个完全显性基因的选择，表达式简化为 $u = spq$（近似），有：

$$qp = \frac{u}{s} \qquad \text{（近似）}$$

$$\text{或} \qquad H = \frac{2u}{s}\text{（近似）} \qquad [2.15]$$

其中 H 是杂合体的频率，u 是向显性等位基因突变的突变率。如果突变基因稀少，H 就很接近群体中突变表型的频率。

例 2.3

如果认为平衡状态是可以应用的，那么我们可以用它来获得选择系数已知的显性畸形体的突变率估计值。在 Haldane（1949）描述的一些人类样本中有丹麦研究的显性侏儒症（软骨营养障碍）的情形。侏儒的频率被估计为 10.7×10^{-5}，他们的适合度（$1 - s$）被估计为 0.196。适合度的估计值是在侏儒产生的孩子数目与他们正常同胞所生孩子数目的比较中获得的。根据等式 [2.15]，推算出突变率为 4.3×10^{-5}。虽然由于侏儒出生前的死亡之故，在频率的估计值中可能存在严重的误差，但突变率几乎无疑地被估计在恰当的数量级之内。用这种方法不能可靠地估计向隐性突变的突变率，因为估计值对小的偏离平衡很敏感（有关人类突变率的更多方面，见 Stern，1973）。

在突变和选择联合作用下的平衡基因频率的这些表达式表明，平衡时基因频率可有任何值，这取决于突变率和选择系数的相对大小。但如果突变率具有普遍被接受的数量级，即 10^{-5} 或其左右，那么只需不利于突变基因的一个温和选择就能把它保持在一个很低的平衡频率上。例如，如果一个基因以 10^{-5} 的速率突变，那么一个 10% 的选择劣势就足以把隐性纯合体的频率保持在万分之一处；而一个 50% 的劣势将把它保持在五万分之一处。因此很明显，自然群体中有害突变体的低频率与从突变和选择的联合作用中预测的大小是相符合的。

现在让我们考察群体处于平衡时的负荷或遗传死亡的比例。正如早期解释的，来自一个隐性基因的负荷是 sq^2。因此，平衡等式［2.12］表明平衡时的负荷是：

$$L = u \qquad [2.16]$$

这样负荷仅取决于突变率，根本不受基因的有害程度的影响。产生这个令人惊奇的结论的原因在于，较有害基因以较低的基因频率达到平衡；因此纯合体较少，并且它们中有较多的死亡。对一个较少有害的基因来说，则有更多的纯合体，且它们中很少有死亡的。因为每一个位点的负荷不依赖于选择系数，所以来自所有位点上的隐性等位基因的总负荷就简单地是突变率的总和$\sum u$。对有害的显性基因，纯合体这样稀少以致于它们可被忽略。因此负荷来自杂合体的死亡，因而 $L = sH$，其中 H 是杂合体的频率。代入等式［2.15］中的杂合体的平衡频率，给出平衡时的负荷为：

$$L = 2u \qquad [2.17]$$

其中 u 是向显性等位基因突变的突变率。而且负荷不受基因的有害程度的影响。比较等式［2.16］和［2.17］引出了另一个问题。为什么来自一个显性基因的负荷应该是来自具有相同突变率的一个隐性基因的两倍？理由是一个突变型纯合体的死亡从群体中淘汰了两个基因，而一个杂合体的死亡只淘汰了一个突变基因。等式［2.17］似乎表明由一个杂合体的死亡引起的一个基因的丢失抵消了由突变引入的两个基因。但并非如此，因为死亡引起的丢失是以每一个体表示的，而突变引起的增量是以每一个配子表示的：每一个体的突变率是 $2u$。来自部分显性等位基因的负荷在 u 和 $2u$ 之间，来自所有位点的总负荷在$\sum u$和$2\sum u$之间，其中 u 是向具有任何显性度的等位基因突变的突变率。如果突变率约为 10^{-5}，那么具有能突变成有害等位基因的10 000个位点的生物体将具有 10% 和 20% 之间的总负荷；也就是说，10 个合子中大约有 1 个或 2 个合子将由于突变而死亡。

低频率的隐性基因对选择起反应非常缓慢的事实，使得稀有隐性基因在实际群体中很难处在它们的平衡频率。除非环境条件在长时间内保持罕见的稳定，否则选择系数的改变很可能比选择能调整频率以适应每一个新的平衡值还要快。当然，这不适用于在所有条件下的致死基因。关于现今人类群体不处于平衡的认识，还有另一个原因，它也适用于致死基因。现代文明社会减少了地方群的再分化，特别是近亲交配，从而降低了纯合体的频率（这一点将在下一章解释），结果使基因频率和纯合体频率两者都低于它们的平衡值，推想它们目前一定正在向较高的新平衡值缓慢地推进之中。因此，以平衡推测为基础不能作出关于稀有隐性基因的推论，特别是对人类群体。

平衡的改变

通过人工产生的辐射或环境化学制剂能够增加突变率；通过医学治疗或通过驯养能降低选择系数，或者通过优生措施可增加选择系数。从这些变化中可预期获得什么效应呢？无论变化是什么，将有一个新的平衡基因频率，群体将开始移向它。当达到新的平衡时，对受影响个体频率的影响可很容易地从等式［2.13］和［2.15］中看出来；例如，加倍突变率或选择系数减半，最终将使受影响个体的频率加倍。增加突变率的直接影响取决于不利于基因的选择系数——选择系数越低，接近平衡就越缓慢。增加突变率的后果无疑将是有害的（对后果的评估见 Crow，1957）。可是，不管如何改变选择系数，其后果都需要一些评述。

选择的强化有时被提倡为人类群体中的一种优生措施。例 2.2 表明对于一个隐性基因应用这种措施引起值得的频率减少将是何等的极其缓慢。当考察突变时，可以看出其前景甚至更糟糕。突变不仅妨碍选择，而且它为 $s=1$ 的平衡施加了一个限制：频率不能减少到低于它。而且，严重的缺陷已经有一种对它们起作用的相当强烈的自然选择，而外加的人工选择充其量也不过仅能使选择系数 s 等于 1。这样做也许不会把目前的选择系数加到一倍以上，有缺陷个体的出现也不会少于目前数值的一半。

可能在现代条件下自然选择的强度降低应当引起我们更多的关注。次要的遗传缺陷，诸如色盲，推想在过去可能有一些选择性劣势，但现在它对适合度的影响即使有一些，也是很小的。而且，医学治疗消除或降低了不利于各种疾病易感性的选择压，这些疾病至少都有某种程度的遗传成因。自然选择的这种放松表明有关基因的频率将向较高值的新的平衡方向推进。如果这是真实的话，我们定可预料到次要遗传缺陷的发生在将来会增加，并且需要对多种疾病进行医学治疗的人的比例也会增加。看来我们医治好了我们自己，却可能在遥远的将来为我们的后代留下了不适宜的基因种子。

有利于杂合体的选择

关于适合度我们考虑了对部分或完全显性的基因起作用的选择效应；虽然在表2-2中给出了适当的公式，但我们还没有讨论关于适合度是超显性的后果；也就是说，当杂合体比任何一种纯合体都有较高的适合度时选择有什么后果。乍看起来不大可能，选择会有利于两个等位基因的杂合体而不是有利于两个纯合体中的任何一个。可是，有可靠证据表明它的确发生了，虽然在它是怎样普遍的一种情形上意见不一致。首先让我们检查这种形式选择的后果，然后考察它如何起作用。

对一个部分或完全显性的基因起作用的选择，倾向于完全淘汰一个或另一个等位基因，在没有突变的情况下，最终的基因频率是 0 或 1。可是，当选择有利于杂合体时，即使没有突变，基因频率也会朝着一个中间值趋于平衡，使两个等位基因都保持在群体中。其理由如下。一个世代之后基因频率的变化给在表 2-2 中，为：

$$\Delta q = \frac{pq(s_1 p - s_2 q)}{1 - s_1 p^2 - s_2 q^2}$$

平衡条件是 $\Delta q = 0$，当 $s_1 p = s_2 q$ 时可满足这个条件。因此，在这个平衡点上的基因频率是：

$$\frac{p}{q} = \frac{s_2}{s_1} \qquad [2.18]$$

$$\text{或} \qquad q = \frac{s_1}{s_1 + s_2} \qquad [2.19]$$

现在，如果 q 大于其平衡值（但不为1），因此 p 小于该平衡值，那么 $s_1 p$ 将小于 $s_2 q$，Δq 将是负值；也就是说 q 将减少。类似地，如果 q 小于其平衡值（但不为 0），那么它将增加。因此当基因频率为 0 或 1 以外的任何值时，选择改变它趋向于等式 [2.19] 给出的中间平衡点，两个等位基因永久地保持在群体中。以同样的方法群体中可保持同一位点上 3 个或更多的等位基因。而且，所需要的选择力量是较简单的（见 Crow 和 Kimura，1970，p.277）。值得注意的一个平衡特性是，基因频率不取决于杂合体的优势度，而取决于一个

纯合体与另一个纯合体相比较的相对劣势。因此只要杂合体优于两个纯合体时，不管优越性多么微小，就会在某个或多或少的中间基因频率上有一个平衡点。

对适合度由超显性引起的负荷是 $s_1p^2 + s_2q^2$，来自 Δq 表达式的分母。将等式 [2.19] 的 q 平衡值以及 p 的相似值代入，导致平衡时负荷的表达式如下：

$$L = \frac{s_1 s_2}{s_1 + s_2} \qquad [2.20]$$

分别代入从等式 [2.18] 中获得的 $s_2 = s_1 p/q$ 和 $s_1 = s_2 q/p$，可导致：

$$L = s_1 p = s_2 q \qquad [2.21]$$

因此，与由频发突变引起的负荷不一样，负荷取决于选择系数，并且来自所有超显性位点的总负荷不能以任何靠求总和的简单方法来获得。

下面描述杂合体优势的一个实例，然后讨论适合度超显性的可能原因。可是，首先我们必须检查选择对观察到的基因型频率的影响。

选择和哈迪—温伯格测验 产生哈迪—温伯格基因型频率的条件之一是没有选择。因此，我们认为像前一章描述的那样，如果哈迪—温伯格测验给出观察和期望基因型频率之间的不一致，这将揭示选择的作用，并允许估计基因型的相对适合度。可是，由于下列原因，关于选择能够得出的结论是有限的（更全面的论述参见 Prout，1965）。

广泛地说，选择可以采取两种形式：存活力的差异和生殖力的差异。在选择存活力之后和在选择生殖力之前，经常通过计数成体进行测验。即，对在成体群中观察的基因型频率和据同一成体群基因频率取得的期望基因型频率作比较。一个有效的测验要求期望值应当用选择前的基因频率来计算。如果选择仅针对生殖力，它不会引起差异，因此测验将不揭示对生殖力的选择。可是，如果选择针对已经发生了的存活力，那么期望值便是从错误的基因频率中计算出来的。因此，对存活力的选择造成了差异，但观察和期望的基因型频率之间的差异并不揭示哪一种基因型被选择过；而对任一纯合体或两种纯合体的不利选择，随着错误计算期望值，会导致杂合体过量和两种纯合体不足。通过解答习题 2.10 可以明白这种情况如何发生。简要说，误差来源应是用前一世代的基因频率计算的期望值。如果基因频率没有改变，则不会引起误差，就像处于平衡频率时具有杂合体优势的基因那样。

例 2.4

人类的镰刀形红细胞贫血病是一个众所周知的杂合体优势的例子。它作为一个实例特别有用，因为数据允许进行观察和理论的测验。这种疾病由异常的血红蛋白 S 引起。纯合体患一种严重的贫血病，患者很多人因此而死去，这个基因仍然存在于非洲人及其在美国的后裔中，该基因频率太高以致于不能为抵消不利于纯合体选择的突变所解释。高频率的解释是，在对疟疾不断增加抵抗力上，杂合体具有超过正常纯合体的优势（Allison，1954）。从 Allison（1956）给出的数据中可以计算出选择压力，然后人们能够看到这些数据怎样恰当地解释了观察到的基因频率。Allison 对坦桑尼亚一个地区的 287 个婴儿和 654 个成年人的基因型进行了分类。（纯合体可通过典型“镰刀”形的红血球的存在来识别；当血样进行脱氧作用时，通过镰刀形细胞的出现来识别杂合体。）观察到的数目和频率以及根据等式 [1.1] 从它们中计算出的基因频率显示在表 2-3 中。（AA 表示正常的纯合体，

AS表示杂合体，SS表示患贫血病的纯合体。）大多数的分化选择被认为在成年之前发生了，即生存下来的基因型在生殖力方面没有很大差别。因此，婴儿代表选择前的基因型频率，成年人代表选择后的基因型频率，这样，可以从观察到的频率中计算出选择系数。可是，首先注意到如果基因频率处于平衡之中，那么选择后的基因频率与选择前的相同，数据与这个期望值吻合得很好。选择后的每一个基因型的频率除以选择前它的频率就给出该基因型的相对适合度，正如表2-3所示。于是将每一个纯合体的适合度除以杂合体的适合度，纯合体的适合度可用相对于杂合体的适合度来表示。纯合体的适合度是$1-s_1$和$1-s_2$，从中可解出不利于AA的选择系数为0.24，不利于SS的选择系数为0.80，两者都相对于AS。根据等式[2.19]，预期由不利于两个纯合体的这种选择引起的平衡基因频率是$q_S=0.23$，它合理地接近观察值。因此，在不同的存活力中观察到的选择压力确实满意地解释了这个群体中镰刀形细胞基因的频率。

表2-3　人红细胞贫血病的若干参数

	基因型			S基因的频率
	AA	AS	SS	
婴儿数	189	89	9	
成年人数	400	249	5	
婴儿中的频率	0.6585	0.3101	0.0314	0.1864
成年人中的频率	0.6116	0.3807	0.0076	0.1980
相对适合度	0.9288	1.2277	0.2420	
相对AS的适合度	0.7565	1	0.1971	
选择系数	$s_1=0.2435$		$s_2=0.8029$	
	期望 $q_S=\frac{s_1}{s_1+s_2}=0.2327$			

如果选择值相对于正常纯合体而表示，那么它可能更有意义。于是AS的适合度是$1/(1-s_1)=1.32$，SS的适合度是$(1-s_2)/(1-s_1)=0.26$。这样，对疟疾的抗性授予了杂合体32%的优势，当基因频率约为0.2时，这平衡了患贫血病的纯合体中74%的劣势。

适合度超显性的可能原因　现在让我们考察对杂合体有利的选择可能起作用的一些方式。一种方式是通过基因多效性，即具有一种以上表型效应的基因。为了产生适合度的超显性，等位基因必须在相反的方向上影响适合度的两个组分。镰刀形细胞贫血病的杂合体优势就是由这种方式引起的；一个纯合体通过一个组分（贫血病）降低了适合度，而另一个纯合体通过另一个组分（即对疟疾的易感性）降低了适合度。人类中还有一些其他基因，被证实或被怀疑有类似原因引起的杂合体优势。另一个例子是野鼠对抗凝血有毒杀鼠灵的抵抗力（Greaves等，1977）。赋予抗性的基因是显性的，这样杂合体和纯合体都具有抗性。可是，纯合体对维生素K有不断增加的很大需求量，这靠正常饮食满足不了。因此在使用了毒物的地区，一个纯合体靠毒物来进行不利选择，另一个纯合体靠维生素K不足来进行不利选择，这导致了抗性基因的平衡频率，在所研究的地区它大约为0.34。

一个位点可以采用很多方式影响适合度的不同组分。例如，组分可以是生活史的不同阶段、在不同时间由同一个体遇到的不同环境、或在不同地方由不同个体遇到的不同环

境、两个性别、修饰所考察位点的其他位点上基因的不同组合等。产生适合度的超显性条件是等位基因在相反方向上影响组分以及在组分结合起来产生适合度的尺度上存在一定的显性度。最后一个条件的意思是这样的：如果组分被相乘在一起以产生适合度，那么在几何尺度上但不一定在算术尺度上必须具有一定的显性度。举一个简单的例子，小鼠有含两个等位基因的一个假想位点，它可能影响每窝出生的数目以及窝数（表 2-4）。

表 2-4　小鼠两个等位基因的一个位点对窝数及每窝数目的影响

	基因型		
	A_1A_1	A_1A_2	A_2A_2
每窝数目	6	7	8
窝数	8	7	6
总数 = 适合度	48	49	48

适合度是两个组分的乘积，并且适合度存在超显性。在它们单独对组分的影响中，等位基因在算术尺度上无显性，但在几何尺度上具有小的显性度，纯合体的几何平均值是 6.9。以这种方式产生的超显性被称为边际超显性（Wallace，1968），意味着超显性仅出现在表的边缘。

以一种类似方式，连锁位点的配子期不平衡能产生拟超显性。如果两个位点紧密连锁，这样它们似乎是一个位点，如果有利等位基因呈显性，并是相斥连锁，那么杂合体可能优于两个纯合体中的任一个。由连锁引起的拟超显性的可能性使得根据从不同品系杂交中获得的群体观察值来证实单个位点的真实超显性极其困难，因为排除一个紧密连锁但没有识别出的位点存在与否，在形式上是不可能的。很多果蝇（*Drosophila*）种的野生群体有倒位片段上携带不同基因排列的染色体。倒位的杂合体通常在适合度方面优于纯合体（见 Wallace，1968），这种杂合体优势可能由倒位中基因的连锁引起。

最后，超显性能出现在分子水平上。如果一个位点编码一种酶，那么两个等位基因的产物（等位酶）可能有不同的特性，诸如酶活性、热稳定性、或环境因子如温度或 pH 的最适宜值。因此等位酶的混合可能使得杂合体比具有单个等位酶的两个纯合体的任一个具有较多的功能，即对由环境条件引起的酶功能损害较不敏感。或者，如果等位酶的活性不同，那么杂合体的中间活性可能比纯合体的较高或较低活性更有利。分子水平上超显性证据的进一步的细节和讨论，见 Berger（1976）。

我们已经看到有很多方式可以引起适合度的超显性。可是，必须承认已被证明发生的情形事实上是很少的。

多态性

早期我们看到突变和选择之间的平衡满意地解释了低频率有害基因的存在，这些基因导致了稀有的畸形或突变个体的出现。可是，这类基因仅是自然群体中发现的遗传变异的较小部分。有很多基因可引起既不稀有又不是任何方面畸形的变异，这些基因的存在很难用对突变的选择所造成的简单平衡来解释。在第 1 章中用作例子的血型基因就是这种情形。更显著的例子是在很多物种，尤其是在昆虫类、蜗牛类和鱼类中发现的颜色变异。

当一个位点等位变异频率太高以致于不能为选择—突变平衡所解释时，我们就说这个位点是多态性的，群体在那个位点上表现出多态性。实用上，多态性位点通常被定义为最共同的等位基因频率低于 0.99 的位点。否则，位点就是单态的；也就是说，对一个等位基因而言，所有个体可被视为是纯合的。

经常用来定量描述由多态性位点引起的遗传变异数量的度量是杂合性 H，简单地说，它就是杂合体的频率。H 可以观察值，或者以从观察的基因频率中计算出的期望值来表示。如果群体不处于哈迪—温伯格平衡，则这两种值将是不同的。当大量位点在一起评定时，对所有位点平均的变异数量可以平均杂合性（$\overline{H}$）表示。杂合性可被看作是在随机选择的任一位点上预期的杂合体频率，或者是任一个体预期是杂合的位点比例。有时用于变异性的另一种度量是在被测验位点内多态性位点的比例。可是，这种度量有不利之处：它对几乎对变异没有贡献的稀有等位基因提供了不合适的权重；它取决于样本大小，因为较小的样本将检测出很少的具有稀有等位基因的位点；选择用来定义多态性的基因频率，在某种程度上是任意的。

几乎在所有的自然群体中都可发现多态性，在遗传组织的所有水平上，从 DNA 序列到主要形态性状，都出现多态性。不同种类性状的变异数量有所不同。造成可见形态变异的等位基因是稀有的。血型和等位酶变异要普遍得多，经常是由处于中间频率的等位基因引起。等位酶的杂合性在分类群之间有所变化。对大量哺乳动物物种进行平均，$\overline{H}$ 是 0.04。无脊椎动物的比脊椎动物的高，但极少超过 0.3（Kimura，1983，p.254）。等位酶多态性也因酶功能的不同而有所不同。在对 9 个果蝇属（*Drosophila*）物种的一项调查中，12 种酶的一个功能群，其 $\overline{H}$ 是 0.10，而 13 种酶的另一个功能群，其 $\overline{H}$ 是 0.23（Latter，1981）。小卫星（VNTR）位点的每一个都具有复等位基因，且是高度杂合的，$\overline{H}$ 几乎为 1。当考虑其中几个位点上的变异时，除同卵双生这一唯一例外，个体可通过它们的基因型来作唯一的识别。这就是用于法庭鉴定的遗传“指纹图谱”。在 DNA 序列变异的水平上，杂合性是每千个碱基对（kb）中若干核苷酸差异的级数。在果蝇（*Drosophila*）被研究的 20 余对基因中，它平均是每 kb 为 3 个（Aquadro 和 Begun，1993）。杂合性变化在很大程度上取决于位点的位置，在限制性重组区域杂合性很低。核苷酸杂合性也随着一种多态性是否造成一种蛋白质中一种氨基酸的改变而有所不同，同义密码子的核苷酸杂合性比异义密码子的要大。

多态性是数量性状变异的来源，是本书的主题，因此多态性及其原因对以后各章具有根本的重要性。对几个可能的原因在下面作一概述。

1. *杂合体优势*　在前一部分我们看到适合度的超显性保持中间水平的平衡基因频率，并且有很多方式可以引起适合度的超显性。因此，杂合体优势是对多态性的一种有吸引力的解释。充分证实这种情况是困难的，因为杂合体优势那么小，以致于实践中不能完全被检测，但却足以保持抑痘粥态性。可是，如果大量位点通过杂合体优势来保持多态性，则会遭受到与遗传负荷和近交效应相关联的困难。另外，已发现单倍体生物的隔离群与二倍体生物具有同样多的等位酶变异性，这个事实强烈反对杂合体优势作为多态性的一种重要原因。这个问题的进一步讨论见 Lewontin（1974），Berger（1976），Wills（1978）和 Kimura（1983）。

2. *取决于频率的选择*　不管表型是什么，具有一种稀有表型的本身可能就是一种优势。于是选择的方向取决于基因频率：低频率的一个等位基因产生稀有表型，这个基因是有利的，但高频率的同一等位基因是不利于选择的。这导致一个稳定的平衡基因频率，因此导致了一种平衡的多态性。已知有很多选择取决于频率的例子。产生一个稀有的自交不育等位基因的花粉粒具有更多的机会使一个胚珠受精，因为相同的等位基因很少出现在其他植物的柱头中。鸟和鱼一旦有了机会，它们会不成比例地获取较多较普遍类型的食物

(例如，见 Allen，1975)，这样，它们便对多态性的被捕食动物（例如蜗牛）施加了取决于频率的选择，给具有一种稀有着色类型的个体提供了优势（见 Clarke，1969)。一般而言，在捕食动物和被捕食动物之间以及病原体和寄主之间的关系中，进攻以及防卫的特殊方法的发展似乎很可能导致取决于频率的选择。Ayala 和 Campbell（1974）以及 Clarke（1979）综述了取决于频率的选择。

3. 杂合的环境 一个群体中的个体所经历的环境不是稳定不变的；它在各个地方不同并且随时间而变化。如果一个等位基因在一个环境中是有利的，另一个等位基因在一个不同环境中是有利的，那么杂合体不必处于平均优势也会产生稳定的多态性。选择可被认为倾向于使不同个体适应不同的环境。情形是复杂的，因为结果取决于很多因素，诸如显性关系，个体是否选择在其适应的环境中繁殖，交配在来自同一环境的个体间是有选择的还是随机的，总群体中居住于每一个不同环境的比例是多少，每一个个体仅经历一种环境（“粗疏纹理”环境）还是一种以上的环境（“细致纹理”环境），以及环境的异质性是空间上的还是时间上的。在杂合环境中一种相对简单形式的选择导致了一种渐变群。这是一个地方和另一个地方之间的一种基因频率梯度，一个等位基因以高频率位于渐变群的一端，另一个等位基因以低频率位于渐变群的另一端。或者在已知的很多情形中，都认为渐变群是通过在一个地区选择一个有利等位基因而在另一个地区选择另一个等位基因来保持，并具有有限数量的迁移，这种迁移只允许来自渐变群的邻近部分的个体之间进行交配。在两端中相反方向的选择和在渐变群中上下起伏的“基因流动”保持了多态性。Felsenstein（1976）综述了杂合环境在保持多态性中的作用。杂合环境具有重要作用的证据是从对 243 个物种的多态性和生态学的调查中获得的（Nevo，1978)。可是，理论上的考虑表明保持多态性的条件是相当受限制的（见 Hoekstra，Bijlsma 和 Dolman，1985)。

4. 过渡 目前看到的多态性可能是一个等位基因被另一个等位基因进化替代过程中的过渡阶段，借助过去某种环境变化，它变得更有利。可是，这似乎不可能是巨大比例的多态性的解释。

5. 中性突变 所有上面的机制都包含了将选择作为对多态性起作用的一种力量。然而，选择系数实际上很小，以致于突变等位基因对适合度是有效中性的，这是可能的。于是，突变和偶然丢失之间的平衡引起了多态性，这取决于群体大小。在描述了群体大小的效应之后，这种理论将在第 4 章中描述。

由于几个原因，获得一种多态性由任一选择性平衡来保持的证据是困难的。可能要从选择系数的直接观察中来寻找选择起作用的证据；但要检测低于几个百分率的选择系数则需要巨大的样本。基因频率的变化可以提供选择的证据；但如果多态性是稳定的，则基因频率在世代间就不发生变化，并且观察到的任何变化只是一个世代内的一种暂时变化。可以用人工方式改变基因频率来看看自然选择能否使其恢复它的初始值；但对于这类干扰实验，必需确信实验条件准确地模仿了自然环境。即使无论通过什么方法检测出选择，也很难确定它的靶子是所关注的多态位点还是与多态位点处于连锁不平衡的某个其他位点。

尽管存在这些困难，仍有 150 多种关于选择在自然群体中起作用的情形被证实（Endler，1986)。例如，在有利于果蝇 *Adh* 等位酶的快/慢多态性平衡选择的过程中，存在直接和间接的证据（Clarke，1975；Briscoe，Robertson 和 Malpica，1975；Hudson，Kreitman 和 Aguade，1987)。可是，大多数单位点情形包含了具有清晰表型的多态性，例如，蜗牛的

壳色和带型、体色、对化学药品和病原体的抗性等。选择生化多态性的例子不那么普遍。

习 题

2.1 稀有的白花植株出现在翠雀花（*Delphinium*）种群中，这个物种正常具有深蓝色花朵。在落基山脉的一个地区，白花植株的频率是 7.4×10^{-4}。发现白花植株平均每株结143粒种子，而蓝花植株结229粒，种子产量的减少是由传粉者的辨别所引起，这些传粉者是大黄蜂和蜂鸟。假定白花是由一单个隐性基因引起的，并且群体处于平衡，需要多大突变率来抵消选择？

数据来自 Waser，N.M.& Price，M.V.（1981）*Evolution*，**35**，376～390。

[解答序号：12]

2.2 如果习题2.1中白花是由一完全显性基因引起的，那么将需要多大的突变率来保持平衡？

[解答序号：22]

2.3 如果一个等位基因 *A* 突变成 *a* 的频率是1/10 000，回复突变的频率是1/100 000，并且如果3个基因型有相等的适合度，那么在一个随机交配群体中平衡时的基因型频率将是多大？

[解答序号：32]

2.4 关于习题2.3。两个方向上的突变率都加倍的后果是什么样？

[解答序号：42]

2.5 医学治疗对几种严重的常染色体隐性疾病是有效的或将是有效的。如果治疗可使这种疾病患者平均而言具有正常人所生孩子数目的一半孩子，而没有治疗他们就没有孩子，这种长期选择的后果是什么？假定目前的频率是突变对选择的平衡，在长期中将达到一个新的平衡，并且其他条件不变。

[解答序号：52]

2.6 囊性纤维变性是一种常染色体隐性的人类疾病，在高加索人中间2 500个活婴中大约有1人发生此种病。如果突变率加倍，那么其下一世代中后果将是什么？假定现在的频率是突变对选择平衡，回复突变可以忽略不计，并且患病的个体没有孩子。将你的结果表达为发病率增加的百分率以及每百万婴儿中附加病例的数目。

[解答序号：62]

2.7 一个粗心的原种果蝇（*Drosophila*）饲养者使显性常染色体突变型的一群原种果蝇被野生型果蝇给污染了。最初，突变型的所有果蝇都是纯合的，但10个世代之后在群体中发现了一些野生型果蝇。于是采取了预防措施以防止进一步污染。假定我们作出下面的假设：(i) 在每一世代中，1%的果蝇是污染蝇，(ii) 所有污染蝇都是纯合的野生型，(iii) 突变型和野生型果蝇有相等的适合度。借助这些假设，(1) 最后一次污染之后的世代中野生型果蝇的比例将是多少？(2) 在具有突变表型的果蝇之间杂合体的比例又是多大？

[解答序号：72]

2.8 在习题1.7中描述的果蝇的两个紧密连锁的隐性基因可看作为等位基因。建立了两个群体，*so* 的起始基因频率在一个群体中为0.2，在另一个群体中为0.8。经7个世

代的随机交配之后，*so* 的基因频率在两个群体中都接近 0.35。关于选择作用，这能告诉我们什么？

［解答序号：82］

2.9　使野生大鼠对抗凝血剂有毒杀鼠灵有抗性的基因表现出杂合体优势，因为具有抗性的基因纯合的大鼠患有维生素 K 缺乏症。杂合体对毒物有抗性，不患维生素 K 缺乏症。死于维生素 K 缺乏症的有抗性的纯合体的比例被估计为 63%。当一个地区施用毒物时，易感的纯合体并不是全部被杀死。在用毒物连续处理下的一个群体达到了平衡，其抗性基因频率为 0.34。由于抗性基因和放毒的结果，这个群体中所有大鼠的死亡百分率是多少？

数据来自 Greaves，J.H.，等（1977）*Genet. Res.*，**30**，257～263。

［解答序号：92］

2.10　假定两个突变基因用于果蝇中关于选择的一个分类实验。在两种情形中，杂合体可与纯合体相区别，但与适合度有关的基因是隐性的（这些不是已知的基因）。对基因（a），两个性别的突变纯合体相对于其余基因型其繁殖力降低了 50%，但具有未受损害的存活力。对基因（b），突变纯合体完全受精，但相对于其余基因型，两个性别的成年前死亡率增加了 50%。在两种情形下，亲本群体由 30♂♂ + 30♀♀纯合野生型和 20♂♂ + 20♀♀纯合突变型组成。在后裔中将发现多大的基因型频率？它们如何与基于后裔中观察到的基因频率的哈迪—温伯格期望值进行比较？关于选择，从后裔的频率中可得出什么结论？为什么在两种情形中 Δq 不同？

［解答序号：102］

2.11　假定由一个隐性基因控制的一个性连锁性状其基因型处于哈迪—温伯格频率。然后一个育种者淘汰两个性别中所有患病的个体。采用起始基因频率的形式来推导由一个世代选择引起的基因频率变化的表达式。

［解答序号：112］

2.12　当选择被速率为 u 的突变所平衡时，一个有害的性连锁隐性基因的近似平衡基因频率是多大？在英国的一项调查中发现人类 X 连锁的肌肉萎缩症其发生率为100 000个男性中有 32.6 个人患有此症。从散发病例数目中估计出突变率为 10.5×10^{-5}。当这种病的患者不生殖并且携带者具有正常的存活力和生殖力时，这些估计值与处于平衡的群体的期望值符合吗？

数据来自（Gardner－Medwin，D.（1970）*J. Med. Genet.*，**7**，334～337。

［解答序号：122］

2.13　牛的很多品种中红毛色是由一个常染色体隐性基因引起的，显性表型是黑色。假定 1%的红毛色犊牛出生在主要是黑毛色的品种中，并假定期望淘汰红毛色基因。假定起始群体的基因型处于哈迪—温伯格比例中，那么采用下列的备择选择方法经过两个世代之后，将有多大比例的红毛色犊牛？（1）红毛色牛不用来繁殖。（2）除淘汰所有红毛色牛以外，用于繁殖的所有黑毛色公牛首先通过 6 个后裔进行测验，这 6 个后裔每一个都来自已知是杂合体的母牛。在这次测验中产生一头或多头红毛色犊牛的任何公牛都被淘汰。用于繁殖的母牛不进行测验。

［解答序号：132］

第3章　小群体：Ⅰ．简化条件下基因频率的变化

现在我们必须考虑最后一种能改变基因频率的作用。这就是离散过程，它区别于系统过程，在于方向是随机的，只能在数量上加以预测。为了从以前的讨论中剔除这个过程，我们总是设定一个“大”群体，我们已经见到在一个大群体中基因频率是内在稳定的。也就是说，如果没有迁移、突变或选择，基因和基因型频率在世代间保持不变。这种稳定性的特性在小群体中不能保持，基因频率因配子的抽样作用而经受随机的波动。把基因传递给下一代的配子携带亲代基因的一个样本，若样本不大，基因频率就很可能在一代和下一世代间发生变化。这种基因频率的随机变化就是离散过程。

在这一章及下一章中我们将涉及离散过程对基因频率的影响。如果关于基因频率进行的推导显得相当远离现实，那么应当记住群体中有关遗传决定的任何性状的特性都取决于基因频率。因此，结论与以后各章要论述的数量性状充分有关。

概括地说，离散过程有4种后果，将在这一章中对它们进行解释和数量化。实际上这些不是不同的后果，相反，它们是可以看到后果的不同方法。它们是：

1．随机漂变　基因频率的随机变化称为随机漂变。如果跟踪任一小群体的基因频率，就会看见基因频率在世代间无规律地变化，毫无回复到初值的倾向。

2．亚群体间的分化　不同亚群体间独立发生的随机漂变导致亚群体间的遗传分化。自然界中栖居在一个大区域中的动物很少组成一个单一的大群体，因为同一区域的栖居者间的交配更经常发生。因此，自然群体或多或少地再分为若干地区群或亚群体，如果亚群体中个体数很少时，亚群体的基因频率就会变得不同。家养或实验室群体也往往以同样方式被细分，例如，分成许多畜群或品系——其中的亚群体以及遗传分化往往更显著。

3．亚群体内的一致性　每一个亚群体内的遗传变异逐步降低，而个体在基因型上变得越来越相似。这种遗传一致性是生物研究的许多领域广泛使用实验动物近交系的原因（一个近交系相当于一个能延续数代的小群体）。

4．递增的纯合性　以牺牲杂合体为代价纯合体频率增加。这种情况加上有害等位基因表现隐性的一般倾向，是近交几乎总是导致生殖力和生活力发生退化的遗传基础。

有两种不同的方法来看待离散过程并推导其后果。一种方法是将它作为一个抽样过程来看待，并用抽样方差的形式来描述它。另一种方法是将其作为一个近交过程来看待，用近亲个体间交配导致的基因型变化来描述它。在这两种方法中，第一种方法对描述过程如何进行可能较简单，但第二种方法为后果的数量化提供了一个更为便利的手段。这里进行的计划是首先从抽样的观点来描述离散过程的一般性质。将显示这4种后果是如何产生的。然后我们重新从近交的观点来研究这一过程，并揭示这两种观点是如何彼此关联的。在所有讨论中，我们只注意最简单的可能情况，不考虑迁移、突变和选择。这样我们就可以看到，在没有影响基因频率的其他因素存在的情况下，小群体中会发生什么情况。在下一章里，我们将剔除限制性的简化条件，将结论扩展到更加现实的情况中去，在第5章

里，我们将考虑系谱群体以及由规则的极近亲交配系统所保持的很小群体等特例。

理想群体

为了将离散过程简化为最简形式，我们设想如下理想群体。我们假定最初有一个大群体，其中交配是随机的，然后这个群体细分为许多亚群体。在自然条件下由于地理或生态的原因会形成这些亚群体，在家养或实验室群体中由于控制其繁殖，也会形成亚群体。最初的随机交配群体被称为*基础群体*，亚群体则被称为*系群*。所有的系群加在一起形成整个群体，而每个系群就是一个“小群体”，其中的基因频率受离散过程的影响。讨论一单个位点时，除非把它作为许多系群之一来考察，否则我们不能恰当地了解一个系群中将发生什么情况。但是，在大量系群中，只要它们都起始于相同的基因频率，在一个位点上的基因所发生的变化，同样也发生在一个系群中大量位点的那些基因上。因此这个过程的一些结果同样适用于一单个系群，只要我们考察其中的许多个位点。

专为理想群体指定的简化条件如下：

1. 交配限制在同一系群的成员之间。系群间没有基因传递，因此在这个意义上各系群是隔离的，也就是说没有迁移发生。

2. 世代区别明确，没有重叠。

3. 每个系群参加繁殖的个体数对所有系群和所有世代都是相同的，繁殖个体是那些能将基因传递给下一代的个体。

4. 各系群内交配是随机的，包括随机数量的自体受精。

5. 任何阶段都没有选择。

6. 不考虑突变。

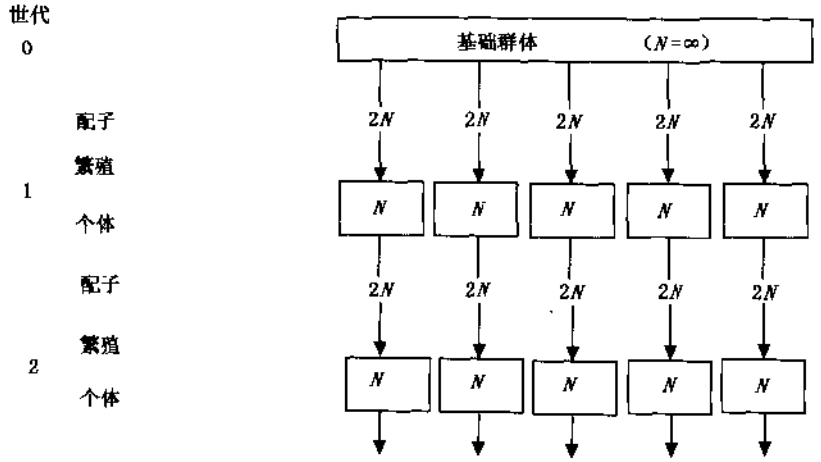


图 3-1 一个单一大群体——基础群体细分成许多亚群体或系群的图解表示

图 3-1 以图解表示出了这些条件所指的情况，并可以这样描述：所有参加繁殖的个体对形成合子的配子库有等量的贡献。配子的结合是绝对随机的。在大量的潜在合子中，只有有限数目的合子成活下来成为下一代的繁殖个体，这就是由配子传递发生的基因抽样阶

段。合子的存活是随机的，因此亲本对下一代的贡献并不一致，依其子代存活的机会而变化。因为群体大小在世代间稳定不变，所以达到繁殖年龄的子代平均数为每个亲本一个后代，或每一对亲本两个后代。在这种方案中，抽样看作是一个单一事件，是大量配子到少量繁殖后代的减少过程。数量的减少可能发生在几个阶段。假设每个阶段的抽样是随机的，则这对从繁殖后代的最终数目中推导出来的理论结果不会产生什么差异。然而，如果在数量减少过程完成之前的阶段对一个群体计数，则观察到的结果将受影响。

用于理想群体的有关符号如下：

N = 在每个系群和世代中繁殖个体的数目，即群体大小。

t = 时间，以世代计，从处于 t_0 的基础群体开始。

q = 一个位点上某一个等位基因的频率。

$p = 1 - q$ = 该位点上所有其他等位基因的频率。q 和 p 指的是在任一系群中的频率；$\bar{q}$ 和 $\bar{p}$ 指的是整个群体的频率，是 q 和 p 的平均值；q_0 和 p_0 是基础群体的频率。

既然所有倾向于改变基因频率的系统过程都已排除，那么在任何阶段所有系群之间的平均基因频率都一定与起始频率相同。因此 $\bar{q} = q_0$，两者在本章中可互换使用。

很明显，为理想群体指定的条件在实际群体中并不成立。可是，通过用“有效”群体大小 N_e 来替代群体大小 N 这一简单方法，本章中获得的结论就可适用于实际群体，有效群体大小 N_e 这个概念将在下一章引入。

抽 样

基因频率的方差

由抽样产生的基因频率的变化其方向是不可预测的，在这个意义上它是随机的。但其大小可以根据变化的方差来预测。考虑来自基础群体的系群的形成。每一个系群由基础群体中抽出的 N 个个体的一个样本构成。由于每一个体在一个位点上携带两个基因，群体的细分代表从基础群体随机抽取的一系列样本，它们每一个各含 $2N$ 个基因。这些样本的基因频率的平均值等于基础群体的基因频率，即 q_0，并围绕这个平均值以方差 $p_0q_0/2N$ 分布，$p_0q_0/2N$ 是样本均值的二项式方差，这种情形中样本大小为 $2N$。这个方差就是 q_1 的方差，q_1 是一代后不同系群的基因频率。因为起始基因频率 q_0 对所有系群都是一样的，因而这一方差也是 $(q_1 - q_0)$，即基因频率变化的方差。这样，由一个世代中抽样导致的基因频率的变化 Δq，可用其方差的形式表示为：

$$\sigma^2_{\Delta q} = \frac{p_0 q_0}{2N} \qquad [3.1]$$

Δq 的这个方差表示由离散过程造成的基因频率变化的大小。它表示任一系群中预期的变化，或一代后在许多系群之间能见到的基因频率的方差。它的效果是基因频率在系群间有所分散；也就是说，尽管从整体上看群体均值没变，但系群在基因频率上有所不同。

下一世代中抽样过程重复进行，但每一系群现在始于一个不同的基因频率，因此第二次抽样导致进一步的离散。变化的方差这次在系群之间不同，因为它分别取决于各系群第一代的基因频率 q_1。通过连续世代连续抽样的效应是各个系群的基因频率作出不规律地变化，各系群逐渐分散，因而造成分化。这些是离散过程开始的两个后果，它们被示例在

图3-2中。如果只有一个小的群体或系群，人们只能看到世代间基因频率不规律变化过程中的随机漂变。图3-2列出了几个系群，从中可见，每个系群都有随机漂变，并且随着系群漂变，系群间也逐步分化。从来自一个不同实验的图3-3中可以更清晰地看出系群之间的分化，该图显示了连续世代中基因频率的分布。

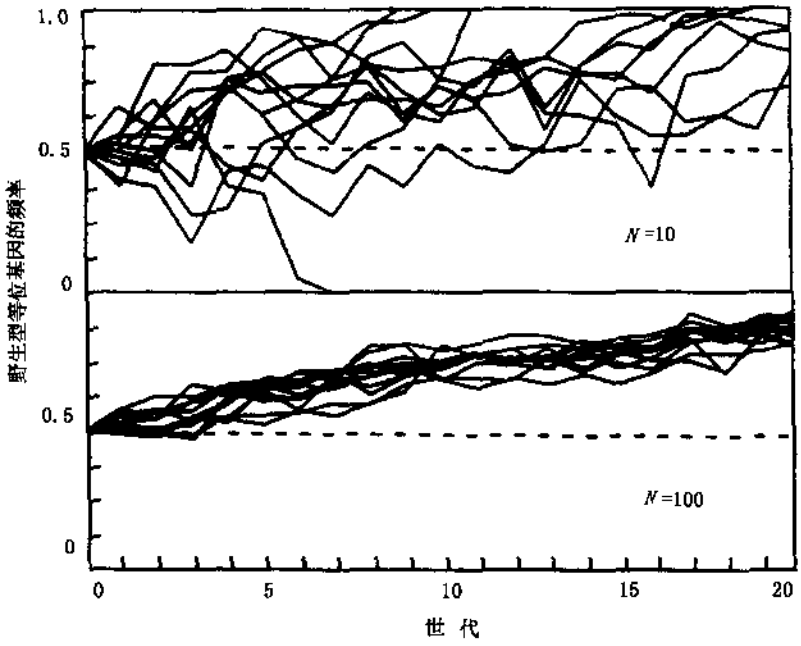


图3-2 赤拟谷盗(*Tribolium*)中一个颜色基因("黑")的随机漂变。杂合体是可识别的，因此通过计数准确估计了基因频率。数字显示了两个群体大小($N=10$和$N=100$)的结果。每个群体大小为12个系群。自然选择对野生型等位基因有利，因而引起其频率全面增加，随机漂变导致各系群围绕平均值产生变异，变异在较小群体中比在较大群体中更明显(根据 Rich, Bell 和 Wilson, 1979)

系群间分化的增加等价于它们之间基因频率方差的增加。在任一世代 t，系群间基因频率的方差 σ_q^2 由下式给出：

$$\sigma_q^2 = p_0 q_0 \left[1 - \left(1 - \frac{1}{2N}\right)^t\right] \qquad [3.2]$$

(该表达式的由来将在后面解释，因为通过考察这一过程的近交更容易理解它)。这里我们会注意到一个后面将需要的事实，根据等式 [3.2]，显然，$\sigma_p^2 = \sigma_q^2$。

对图3-3中基因频率分布的检测表明，分布在形状上不断变化，最终变得非常扁平，所有频率具有相等的可能性（这对于0和1的极限频率不是真实的，下一部分将讨论这一问题)。理论上的考虑表明离散过程有两个阶段。初始阶段，基因频率从初始值开始散开。紧接着是一个平稳阶段，此时基因频率均匀地散布于整个范围内，除两个极限值以外的所有基因频率都有同等的可能性。即使初始基因频率不是0.5，平稳阶段的均匀分布也能达到，虽然这要花费较长的时间。初始阶段持续的世代数为群体大小的小倍数，这取决于初始基因频率。$q_0 = 0.5$ 时，它持续大约 $2N$ 世代，$q_0 = 0.1$ 时，它持续大约 $4N$ 世代

(Kimura，1955)。$q_0 = 0.5$ 和 $q_0 = 0.1$ 的初始阶段基因频率的理论分布示于图 3-4。对于 $q_0 = 0.5$，图 3-3 中的观察分布与理论分布吻合良好。

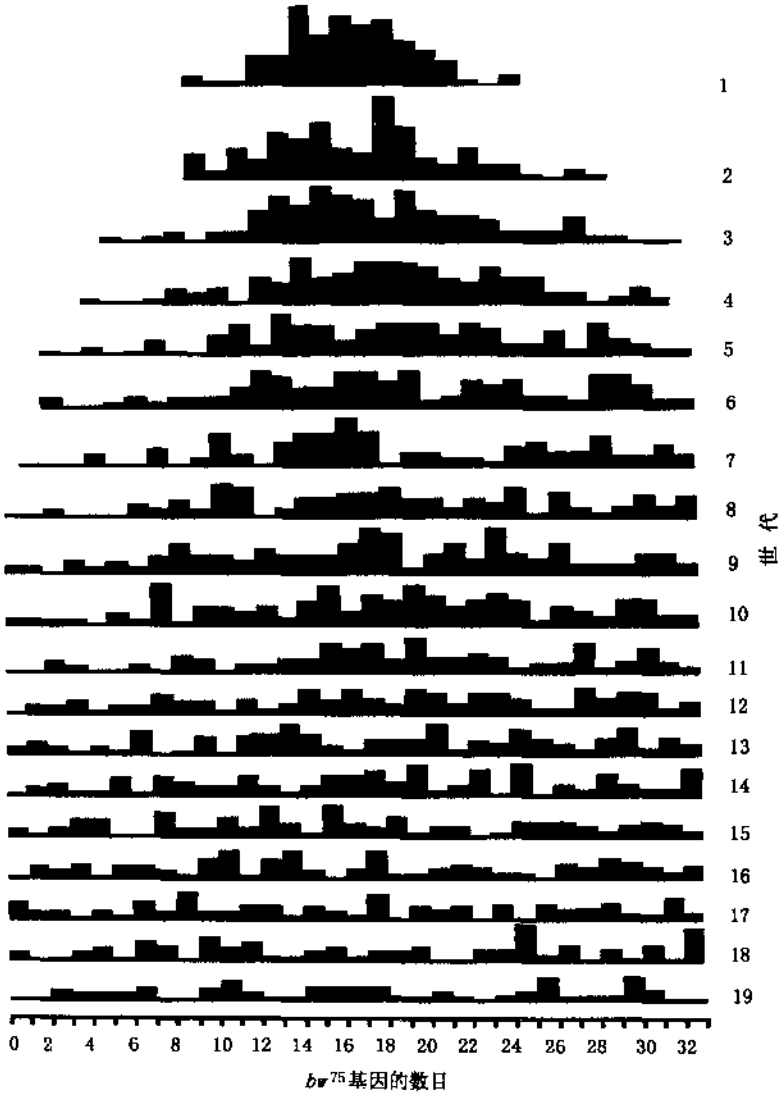


图 3-3　黑腹果蝇（*Drosophila melanogaster*）105 个系群（每一系群有 16 个个体）之间连续 19 个世代中基因频率的分布。基因频率指的是位于 'brown' 位点上的两个等位基因（bw^{75}和 bw），起始频率为 0.5。每一个黑柱的高度显示具有以下面尺度表示的基因频率的系群数，以前固定的系群不包括在内（根据 Buri，1956）

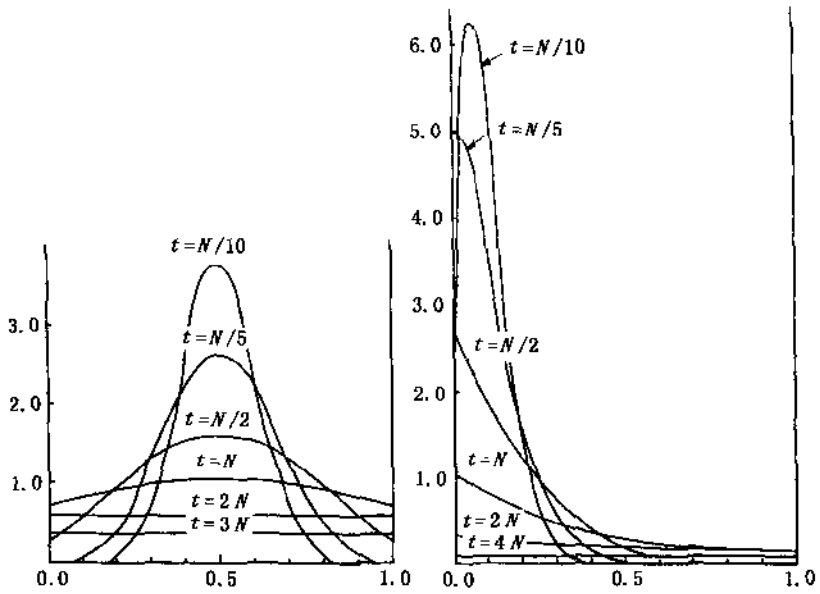


图 3-4　在不同的世代数 t 之后，系群间基因频率的理论分布（以系群的群体大小 N 表示世代 t）。左边图中 $q_0=0.5$，右边图中 $q_0=0.1$。以前固定的系群（见下一部分）不包括在内。任一系群中水平尺度都是基因频率 q。纵轴是概率，其绘制使得曲线下方的面积等于未固定系群的比例（根据 Kimura，1955）

固　定

离散过程引起的系群分散是有极限的。基因频率不可能变化到超出 0 和 1 的两个极限之外，并且迟早每一系群必须达到其中一个极限。而且，极限是“陷阱”或不可回返之点，因为一旦基因频率达到极限 0 或 1，该系群中这个基因频率就不再变化。当一个特定等位基因达到频率 1 时，则称它在该系群中被固定，而当频率达于 0 时，则称为丢失。当一个特定等位基因达到固定时，该系群中不再出现其他等位基因，于是系群被称为是固定的。一个系群被固定后，它的所有个体的这个位点都具有一致的基因型。最终，所有系群以及该系群中所有位点都被固定。于是，一个系群中的个体在遗传上都相同，这就是高度近交系遗传一致性的基础。

当这个过程已经完成，所有系群都被固定后，平均基因频率不变且等于初始基因频率。因此，一个位点上不同等位基因被固定的系群比例等于各等位基因的初始频率。如果基础群体含有两个等位基因 A_1 和 A_2，其频率分别为 p_0 和 q_0，那么 A_1 将以系群的比例 p_0 被固定，A_2 则以剩余比例 q_0 被固定。于是系群间基因频率的方差为 p_0q_0，正如令 t 为无穷大，从等式［3.2］中看到的那样（图 3-3 只显示了刚刚固定或丢失的系群，但没有显示那些早些时候已固定或丢失的系群）。

考虑到遗传一致性的获得，人们想知道固定多久才发生；在一定数目的世代之后，一特定位点被固定的概率是多大，或一个系群中所有位点有多大比例将被固定。正如图 3-3 所示，考虑到离散的渐进特性将看到固定并不立即开始；在任何系群可能达到固定之前，

基因频率的离散必须进行一些时候。要推导达到固定的概率在数学上是复杂的，这里只能给出结论要点。达到平稳阶段后，固定以一种不变的速率继续下去；以前未被固定的系群在每一世代中有 $1/N$ 的比率被固定。在固定的初期阶段之后，在起始基因频率为 q_0 的系群中，预期被固定、丢失或仍分离的比例大致如下（Wright，1952）：

$$\left.\begin{array}{ll} \text{固定：} & q_0 - 3p_0q_0P \\ \text{丢失：} & p_0 - 3p_0q_0P \\ \text{两者都不是：} & 6p_0q_0P \end{array}\right\} \quad \text{其中，} P = \left(1 - \frac{1}{2N}\right)^t \qquad [3.3]$$

基因型频率

基因频率的变化导致基因型频率的变化；因此，小群体中的基因型频率随离散过程引起的基因频率变化而变化。在我们一直考察的理想群体中，每一系群内的交配是随机的。因此，任一系群中基因型频率是相当于该系群上一世代的基因频率的哈迪—温伯格频率（事实上，与系群内哈迪—温伯格频率有小的离差，这将在本章结尾时解释，但现在可以忽略它）。随着系群在基因频率上的漂变分离，它们的基因型频率也随着分化。但分化不是变化的唯一方面：变化的总方向是纯合基因型趋于增加，杂合基因型趋于减少。其原因在于基因频率从中间值趋向极端值的分散。杂合体在中间基因频率时最多（见图 1-1），所以基因频率趋于极端的漂变，平均而言导致杂合体频率下降。它还导致系群中有相同基因型的个体比例增加，从而导致系群内遗传一致性的增加。

整个群体的基因型频率，采用下面的方法可从有关基因频率的方差知识中推导得出。如果一个等位基因在一特定系群中频率为 q，则该等位基因的纯合体在该系群中的频率为 q^2。因此，这些纯合体在整个群体中的频率是所有系群 q^2 的平均值。我们将纯合体的平均频率记作 $(\overline{q^2})$。$(\overline{q^2})$ 值可从系群间基因频率的方差中得出，只要从观察值平方的均值中减去均值的平方就可求得一组观察值的方差。因此：

$$\left.\begin{array}{l} \sigma_q^2 = (\overline{q^2}) - (\bar{q})^2 \\ (\overline{q^2}) = (\bar{q})^2 + \sigma_q^2 \end{array}\right\} \qquad [3.4]$$

其中，σ_q^2 为系群间基因频率方差，如同等式［3.2］给出的，$(\bar{q})^2$ 为平均基因频率的平方。因为平均基因频率 $\bar{q}$ 等于初始频率 q_0，所以 $(\bar{q})^2$ 或 q_0^2 就是基础群体中纯合体的初始频率。这样整个群体中一个特定等位基因的纯合体频率上升，并且总是超过起始基因频率，所超过的数量等于系群间基因频率的方差。在一个两等位基因的系统中，对另一个等位基因的情况同样如此，杂合体频率相应降低。注意到从等式［3.2］，$\sigma_p^2 = \sigma_q^2$，因此，我们可得出有两个等位基因的一个位点上的基因型频率如下：

基因型	整个群体中的频率
A_1A_1	$p_0^2 + \sigma_q^2$
A_1A_2	$2p_0q_0 - 2\sigma_q^2$
A_2A_2	$q_0^2 + \sigma_q^2$

[3.5]

这些基因型频率不再是与起始或平均基因频率相当的哈迪—温伯格频率了。基因频率和基因型频率之间的哈迪—温伯格关系，虽然在各个系群内部分别考虑是适用的，但如把各系群合并起来作为一单个群体来看待就不适用了。在把自然群体中基因和基因型频率关联起来时这种事实造成一些困难，因为群体经常或多或少地被细分，而且细分的程度往往是不知道的。

前面关于基因型频率的讨论描述了在许多系群中以一个位点为根据的情况。这同样可适用于一个系群中许多位点的情况；于是，在任一系群或小群体中的变化就是纯合个体位点数增加，而杂合个体位点数相应减少——简言之，以杂合体的减少来换取纯合体的增加。由离散过程引起的基因型频率变化是近交退化现象的遗传基础，第14章对此将有充分的阐述。

现在，我们已经纵览了离散过程的一般性质及其4种主要后果，即随机漂变、亚群体的分化、亚群体内的遗传一致性和纯合基因型频率的全面增加。现在让我们从另一个角度来看这个过程，把它作为一个近交过程来看待。不再将纯合体的增加作为基因频率分散的一种后果，我们现在将直接探讨使额外的纯合体上升的方式。

近 交

近交意味着因祖先而彼此有亲缘的个体一起交配。与群体大小有关的群体中个体间亲缘关系的程度，显然可通过考察可能的祖先数目而清楚。在两性生物体的一个群体中，每个个体都有两个亲本、4个祖亲、8个曾祖，等等，依此类推，追溯 t 世代就有 2^t 个祖先。不需追溯很多世代，可为所有现存个体提供不同祖先所需要的个体数，便会超出任何实际群体所能容纳的数目。因此，任何一对个体必然因或远或近的世代中一个或多个共同祖先而彼此有亲缘；并且过去世代中群体越小，它们的共同祖先就越近，或它们的共同祖先的数目就越多。这样，在小群体中随机交配的配偶间的亲缘关系要比在大群体中的更密切。这就是为什么小群体的性质可以作为近交的后果来处理的理由。

有一个共同祖先的两个个体的重要后果，是它们可能都携带共同祖先所有基因中某一基因的复制品；而且，如果它们交配，就能将这些复制品传给它们的后代。因此，近交个体——也就是说，由近交产生的后代——可能在一个位点上携带两个基因，而这两个基因是上一代中一个相同基因的复制品。考虑这种近交的后果看出在等位基因之间有两类同一性和两类纯合体。迄今为止我们已经考虑的这类同一性是一种功能上的同一性。如果两个基因不能根据其表型效应或任何其他的功能标准来加以辨别，那么这两个基因被认为是相同的等位基因。在通常的意义上，携带一对这样基因的个体就是一个纯合体。新的同一性类别是复制的来源相同。如果两个基因都来自上代中一个单基因的复制，则可以称之为**血统同一**，或简单地称为**同一**。不是同一的两个基因在血统上是**独立的**。同一基因的纯合体可以称作**同一纯合体**。使用的其他术语是：**同合**用来描述同一纯合体，**异合**用来描述已知不是同合的纯合体。正是同一纯合体的产生才引起近交的一种后果——纯合体的增加。

通过配偶间亲缘关系的程度，血统同一为离散过程的度量提供了基础。这个度量就是**近交系数**，它是一个个体中任一位点上两个基因是血统同一的概率。它适用于一个个体，表示该个体的亲本间亲缘关系的程度。如果任一世代的亲本随机交配，那么子代的近交系数就是从亲代随机获取的两个配子在一个位点上携带同一基因的概率。这是所有子代的平

均近交系数。不同家系的个体将具有不同的近交系数，因为对于随机交配，一些亲本对的亲缘关系要比另一些的更密切。可是，我们关注的正是用平均近交系数作为离散过程的一种度量。近交系数通常用符号 F 表示。

近交系数所表示的亲缘关系的程度，主要是被讨论群体与某个指定的或隐含的基础群体之间的一种比较。如没有这个参考点，它是没有意义的，下面的考察将阐明这一点。鉴于在任何含量不是无穷大的群体内独立祖先在数量上的有限性，如果追溯得足够久远，则会发现群体中一个位点上的现有基因都是血统同一的。因此，近交系数只在我们指定的过去一段时间中才有意义，此前的祖先不去追溯，而且此时群体内现存的全部基因都认为是独立的，即不是血统同一的。这一点就是基础群体，根据它的定义，其近交系数为零。一个后继世代的近交系数表示自基础群体以来已经发生的离散过程的数量，以及现有个体间的亲缘关系程度与基础群体中个体间的亲缘关系程度进行的比较。关于础群体并不总是说得清楚，但总是含蓄的。例如，我们可以谈到一个细分为许多系群的群体的近交系数。亲缘关系的比较是在一个系群的个体与从整个群体中随机抽取的个体之间的比较。所暗含的基础群体是所有系群起源的假想群体。

理想群体中的近交

让我们现在回到理想群体，从基础群体开始，其子代构成 1 世代，来推导连续世代中的近交系数。这种情形可通过设想一种雌雄同体的海洋生物来理解，这种海洋生物能自体受精、把卵子和精子排在海水中。设有 N 个个体，每个个体产生同样数量的配子，它们随机结合。基础群体中一个位点上的所有基因必须认为是不同一的；因此，只考察一个位点时，基础群体产生 $2N$ 种不同种类的配子，数目相同，在 A 位点携带 A_1、A_2、A_3、…等基因。任何一类配子都携带同一基因；不同类的配子则携带独立来源的基因。随机抽取的一对配子携带同一基因的概率是多少呢？这就是 1 世代的近交系数。任一配子都有$\frac{1}{2N}$的机会和同类的另一配子结合，因此$\frac{1}{2N}$是结合的配子携带同一基因的概率，因而是子代的近交系数。现在考察第二代。现在有两种方式可以产生同一纯合子，一是来自基因的新的复制，另一种是来自以前的复制。一个合子中结合在一起的新复制的基因概率仍是$\frac{1}{2N}$。合子的剩余部分，$1-\frac{1}{2N}$，携带的基因以 1 世代为起源是独立的，以 0 世代为起源也许是同一的。它们在 0 世代中为同一起源的概率就是我们已推导出的 1 世代的近交系数。这样，2 世代中同一纯合子的总概率为：

$$F_2=\frac{1}{2N}+\left(1-\frac{1}{2N}\right)F_1$$

其中，F_1 和 F_2 分别代表 1 世代和 2 世代的近交系数。同样的推论适用于以后的世代，因此，一般地，t 世代中个体的近交系数为：

$$F_t=\frac{1}{2N}+\left(1-\frac{1}{2N}\right)F_{t-1} \qquad [3.6]$$

这样近交系数由两部分组成：一个“增量”$\frac{1}{2N}$，归因于新的近交；一个“剩余量”，归因

于以前的近交并具有以前世代的近交系数。在理想群体中，“新的近交”是产生于自体受精，它直接把上一代复制的基因联合在一起。排除自体受精，简单地将复制推回一个世代，这样“新的近交”便将祖父母代复制的基因联合在一起；近交系数会受到影响，但影响不太大，这一点我们以后还会见到。“新的”和“旧的”近交间的区别阐明了一个要点：如果没有“新的近交”，如同在群体大小突然增大时所发生的情况那样，则以前的近交不会消失，而是保持群体大小增大以前的状态。这里我们顺便提一下，因为它在以后还需要提到，并在实践中经常是重要的。

假定我们将“增量”或“新的近交”称作 ΔF，这样：

$$\Delta F = \frac{1}{2N} \qquad [3.7]$$

于是等式［3.6］可重写成这样的形式：

$$F_t = \Delta F + (1 - \Delta F) F_{t-1} \qquad [3.8]$$

进一步的重新整理可使增量 ΔF 的精确意义更明显：

$$\Delta F = \frac{F_t - F_{t-1}}{1 - F_{t-1}} \qquad [3.9]$$

从这样书写的等式中，我们可看到“增量” ΔF 以一种比例增加的形式来度量近交率。这是一个世代中近交系数的增加，相对于为达到完全近交还需走的距离。这个近交率的度量为超越理想群体的有限简化提供了一条便利的途径，这样它为比较不同交配系统的近交效应提供了一种方法。当近交系数以 ΔF 的形式来表示时，等式［3.8］对任一交配系统都是有效的，不局限于理想群体，虽然只在理想群体中才有 $\Delta F = \frac{1}{2N}$。

到目前为止，我们做的不过是将一代中的近交系数与上一代的关联起来。还需要把等式［3.8］延伸返回到基础群体，这样就可以用世代数的形式来表示近交系数。这很容易做到，采用一个符号 P，代表近交系数的余数 $1 - F$，它被称为随机交配指数。用（$1 - P$）代替等式［3.9］中的 F，重新整理导致：

$$\frac{P_t}{P_{t-1}} = 1 - \Delta F \qquad [3.10]$$

因此，随机交配指数在每一世代中以固定比例下降。延伸回到 $t-2$ 代，给出：

$$\frac{P_t}{P_{t-2}} = (1 - \Delta F)^2$$

延伸返回到基础群体则给出：

$$P_t = (1 - \Delta F)^t P_0 \qquad [3.11]$$

其中，P_0 为基础群体的随机交配指数。基础群体被定义为其近交系数为 0，所以随机交配指数为 1。因此任一世代 t 中，以基础群体为依据的近交系数为：

$$F_t = 1 - (1 - \Delta F)^t \qquad [3.12]$$

前面我们用抽样方差的观点描述了离散过程的结果。现在让我们用近交率和近交系数作为过程的度量再去看一看这些结果。严格地说，我们仍应以理想群体为依据，但把这两种观点同等看待，一般还是有效的，除非群体大小在亲代和子代中不同，或系群内存在非随机交配。

基因频率的方差

首先，一个世代中基因频率变化的方差，由等式［3.1］得来，并用近交率的形式表示，成为：

$$\sigma^2_{\Delta q} = \frac{p_0 q_0}{2N} = p_0 q_0 \Delta F \qquad [3.13]$$

书写等式［3.13］的一种等价方式是用一个世代之后的近交系数和基因频率方差的形式来表示。它表明任何数目的世代之后亲缘关系都是相同的，因此 t 代后：

$$\sigma^2_q = p_0 q_0 F_t \qquad [3.14]$$

可以看出等式［3.14］与没有解释就给出的等式［3.2］是等价的。用等式［3.12］和［3.7］的$\left[1-\left(\frac{1}{2N}\right)^t\right]$代替等式［3.14］中的 F_t 便得出等式［3.2］。

对离散过程的度量，以近交为基础比以基因频率方差为基础更有用，因为它们可以同等地适用于任一平均或起始基因频率。这样 ΔF 表示离散率，F 则表示随机漂变的累积效应。

基因型频率

让我们接着来看整个群体中的基因型频率。等式［3.5］中，用基因频率方差表示的基因型频率可用等式［3.14］中的近交系数来重新书写。例如，A_2A_2 的频率为：

$$(\overline{q^2}) = q_0^2 + \sigma^2_q = q_0^2 + p_0 q_0 F$$

用这种方式表示的基因型频率列入表 3-1 的左边栏中。正如以前解释的那样，书写基因型频率的这种方式显示了纯合体如何因杂合体的减少而增加的情况。近交的观点引导我们对血统同一性的识别意味着我们现在可以在 A_1A_1 和 A_2A_2 基因型之间区分两种纯合体：同一纯合体和独立纯合体。合计的两种基因型之间同一纯合体的频率，根据定义就是近交系数 F；而两种基因型之间的商数是与起始基因频率成比例的。因此，p_0F 是 A_1A_1 同一纯合体的频率，q_0F 是 A_2A_2 同一纯合体的频率。剩余的基因型，包括纯合体和杂合体，携带起源独立的基因，因而相当于从整个群体中随机抽取的配子对。因此，它们的频率是哈迪—温伯格频率。这样，由近交观点看，我们得出了表 3-1 右边栏所示的基因型频率。书写基因型频率的这种方式显示了如何把纯合体划分为独立起源的纯合体和同一起源的纯合体。表示基因型频率的两种方法的等价性可由它们在代数学上的同一性来确证。两种方式同样清楚地显示了杂合体是如何在频率上以比例于 $1-F$ 的值而下降。

随机交配指数，先前被定义为 $P = 1 - F$，表示了一个亚群体中杂合体的频率，相当于整个群体随机交配所期望的哈迪—温伯格频率。这一点可由表 3-1 中给出的杂合体的频率看出。假定 H_t 和 H_0 分别为亚群体和随机交配群体中杂合体的频率。那么 $H_0 = 2p_0q_0$，$H_t = 2p_0q_0(1-F) = H_0(1-F)$。因此，世代 t 的随机交配指数为：

$$P_t = 1 - F_t = \frac{H_t}{H_0} \qquad [3.15]$$

表 3-1 具两个等位基因的一个位点上，以近交系数 F 表示的基因型频率

	起始频率	近交引起的变化	起源	
			独立的	同一的
A_1A_1	p_0^2	$+p_0q_0F$	$=p_0^2(1-F)$	$+p_0F$
A_1A_2	$2p_0q_0$	$-2p_0q_0F$	$=2p_0q_0(1-F)$	
A_2A_2	q_0^2	$+p_0q_0F$	$=q_0^2(1-F)$	$+q_0F$

当一个实际群体被抽样时，唯一能显示它是一个亚群体的现象就是杂合体的缺少。于是，观察到的杂合体频率 H，相当于哈迪—温伯格频率$2\bar{p}\,\bar{q}$，给出随机交配指数为 $P=\frac{H}{2\bar{p}\,\bar{q}}$，$\bar{p}$ 和 $\bar{q}$ 为整个样本中观察到的基因频率。可是，需要谨慎的是，这样计算得来的 P 值只能认为是这个样本的描述。实际存在的所有亚群体不大可能都将同等地出现在样本中，例如，除非是这样，否则 P 不能有效地被用来根据等式［3.14］估计亚群体之间的基因频率方差。

系群内的基因型频率 在本章的各个部分中，我们都假定交配在系群内是随机的，因此任一系群内的基因型频率是与该系群中基因频率的哈迪—温伯格频率相称的。可是已经指出，基因型频率实际上与哈迪—温伯格期望值稍有偏离。这种偏离的原因是传递给下一代的基因样本由两个独立样本组成，一个样本是雄性亲本的，另一个样本是雌性亲本的，每个样本中都有 N 个基因。因此平均而言，雄性和雌性亲本的基因频率有所不同。雌雄亲本间基因频率的差异导致子代中杂合体过多；或者换句话说，从平均基因频率计算出来的期望值太低。将适当的配子频率列入表1-7中，可以看到任一系群内杂合体的期望频率为 $H=2pq+\frac{1}{2}\overline{D^2}$，其中 p 和 q 为系群中的平均基因频率，D 为雌雄亲本间基因频率的差异，$\overline{D^2}$为差异平方的平均值。因为$\overline{D}=0$，根据等式［3.4］类推，$\overline{D^2}=\sigma_D^2$。方差 σ_D^2 为两个大小为 N 的二项样本间差异的方差，其大小为 $2pq/N$。因此，系群内杂合体的期望频率为：

$$H=2pq+pq/N=2pq\left(1+\frac{1}{2N}\right) \qquad [3.16]$$

（进一步的细节见 Robertson，1965）。杂合体的过量部分是不重要的，除非样本的亲本数 N 非常小。但它对单个群体的小样本中观察的频率做哈迪—温伯格期望值符合度的检验会有一定的影响。

一个被细分群体的整体中的杂合体总频率有时被用来估计群体历史上的近交量，而不仅仅作为现存细分状态的一种描述。这在例 4.1 中已做了。如果系群能分别识别，并且每一系群中被抽样的亲本数量已知，则可对杂合体的过量部分作校正。用等式［3.16］中 H 代替等式［3.15］中 H_0，可得出：

$$1-F=\frac{H}{2pq\left(1+\frac{1}{2N}\right)} \qquad [3.17]$$

其中，H 为杂合体的观察频率，p 和 q 为总的观察到的基因频率，N 为每一系群中的亲本数。

习 题

求解第3章中习题时，将所有群体作为理想群体来处理。

3.1 鳕鱼有两种类型血红蛋白，这由一个位点上两个等位基因 a 和 b 确定。由挪威海岸采取的鳕鱼样本得到下面3种基因型的频率：

aa	*ab*	*bb*	总数
130	763	1 698	2 591

这些频率是否与从随机交配群体中抽取的样本符合？关于群体的繁殖结构，它们表明了什么？

数据来自 Møller，D.（1968）*Hereditas*，**60**，1～32。

［解答序号：3］

3.2 习题3.1所述的鳕鱼可依其耳石解剖学差异区分为两个明显的族。当将样本分为两个族，称“北冰洋型”和“海岸型”时，获得下列数据：

	aa	*ab*	*bb*	总数
北冰洋型	23	250	946	1 219
海岸型	127	513	752	1 372

这对3.1中的问题进一步阐明了什么？

［解答序号：13］

3.3 如果一个群体在每一世代中通过20对亲本随机交配来保持，那么5代、10代之后其近交系数是多少？

［解答序号：23］

3.4 假定对一个分组实验，给每个学生10对未交配的果蝇（*Drosophila*），它们是从一大群中随机抽选出来的，该群体中电泳变异型以0.3的基因频率出现。然后每个学生随机抽选10对果蝇作为下一世代的亲本来保持他的亚群体。5代后，每个学生通过来自后代的20个果蝇样本的电泳来确定他自己群体中的基因频率。发现的平均基因频率将是多少？假定所有学生都能正确地读懂凝胶电泳结果，那么你预计会发现学生们估计的基因频率间有多少变异？

［解答序号：33］

3.5 如果将习题3.4的实验中由学生统计的3种基因型的数目加在一起，那么基因型的总频率是多少？

［解答序号：43］

3.6 一群小鼠由18个系群组成，全来自同一个基础群体，此后分别配种繁殖。它们在一个常染色体酶位点 *Got*－1 上具有多态性，该位点有两个等位基因 a 和 b。27世代后，根据该位点基因型的电泳对所有系群的小鼠进行分型，并获得如下数据：

aa	*ab*	*bb*	总数
42	76	448	566

由这些数字表明的近交系数是多少？

数据来自 Garnett，I.（1973）　爱丁堡大学博士学位论文。

［解答序号：53］

3.7　假定对一个随机繁殖群体进行抽样，获得一种蛋白变种的基因型频率如下：

aa	*ab*	*bb*
0.34	0.52	0.14

（1）忽略显著性问题，这些频率是否为对基因型起作用的某种形式的选择提供了证据？（2）若样本中的个体为 4 对亲本的后代，则结论将作怎样修改？

［解答序号：63］

3.8　修改等式［3.16］，使之像家畜中通常情况那样，当雄性和雌性亲本数目不同时能适用。

［解答序号：73］

第 4 章　小群体：Ⅱ．较少简化条件

为了简化对离散过程的描述，上一章中我们把注意力局限于理想群体，这样做，我们必须指定许多限制性条件，而在实际群体中很少能满足这些条件。本章的目的是使上一章的结论适合于不再维持强加条件的情形；换句话说，就是要消除较严格的限制，从而使结论更接近现实。限制条件有两类，一类与群体的繁殖结构有关，另一类从所考察的因素中排除了突变、迁移和选择。我们将首先描述偏离理想繁殖结构的离差效应，然后考察当突变、迁移或选择同时起作用时离散过程的结果。

有效群体大小

如果繁殖结构与理想群体所规定的不相符合，那么仍能按照基因频率的方差或近交率来评估离散过程。这可以用相同的通用方法做到，并不涉及新的原理。因此我们将简要地给出结论而不作详细的解释。这个问题的全面综述见 Caballero（1994）。要处理偏离理想繁殖结构的任一特定离差，最方便的方法是以繁殖个体的*有效数目*或*有效群体大小* N_e 的方式来表示。如果个体按照理想群体的方式繁殖，则这就是产生计算抽样方差或近交率的个体数。例如，假定从考虑产生同一纯合体的概率中已计算出一特定繁殖结构的近交率 ΔF。在理想群体中，依等式［3.7］$\Delta F = \frac{1}{2N}$，ΔF 与群体大小 N 有关。有效大小以同样方式与 ΔF 有关，因此从计算的 ΔF 中可以获得有效大小，即 $N_e = \frac{1}{2\Delta F}$。这样，如果有效数目 N_e 代替实际数目 N，前一章中获得的所有结论对任何繁殖结构都是有效的，推导出的公式都可以应用。当繁殖结构已知时，有效数目能够从实际数目中获得，对于最普遍的偏离理想繁殖结构的情况，下面给出了二者之间的关系。精确的表达经常是复杂的，但在大多数情形下，可以使用具有足够准确度的一种简单的近似。重要的是要注意到，在这些关系中，实际数目 N 是繁殖个体的数目，因此它不可能从调查中获得，除非不同年龄组是可辨别的。知道了任何繁殖结构的有效群体大小 N_e，于是可以获得近交率为：

$$\Delta F = \frac{1}{2N_e} \qquad [4.1]$$

借助前一章的公式，从 ΔF 中可以计算近交的任何结果。

极近亲交配的排除

两性生物体中，自体受精当然是不可能的。在人类中同胞交配已被消除，在实验动物和家畜群体的保持过程中也经常有意避免同胞交配。可是，由于下面的原因，极近亲交配的排除对近交率并未造成巨大的差异。极近亲交配的后裔比关系较不密切交配的后裔具有较高的近交系数。因此它们的存在任何时候都会增加群体的平均近交系数。但它们较高的近交不是永久的：随机交配，它们本身很可能与关系较不密切的个体进行交配，因此它们

高于平均的近交不会传递给其后裔。这样，极近亲交配的排除，总的来说降低了平均近交系数，但它对近交累积的速率没有大的影响。排除的效应可用有效数目大致进行数量化，结果如下（Wright，1969，p.212）。

排除自体受精时，

$$N_e = N + \frac{1}{2} \quad \text{（近似）} \qquad [4.2a]$$

因此，借助等式［4.1］有：

$$\Delta F = 1/(2N+1) \quad \text{（近似）} \qquad [4.2b]$$

同胞交配也排除时，

$$N_e = N + 2 \quad \text{（近似）} \qquad [4.3a]$$

和

$$\Delta F = 1/(2N+4) \quad \text{（近似）} \qquad [4.3b]$$

这些近似值在计算 ΔF 时只产生极小的误差，除非 N 很小，如同极度近交的情况；不过需要推导 ΔF 的其他方法，这将在下一章中阐明。

雄性和雌性数目不同

在家养和实验动物中，繁殖个体之间两性经常不等地出现，因为在可能的情况下，使用较少的雄性和较多的雌性比较经济。可是，无论它们的相对数目是多少，两性贡献下一世代的基因是相等的。因此必须分别计算归因于两性的抽样方差。因为抽样方差与数目的倒数成比例，所以有效数目是两性数目的调和平均数的两倍。它是调和平均数的两倍，因为群体大小是 $N = N_m + N_f$，其中 N_m 和 N_f 分别是雄性和雌性的数目。调和平均数是 $1/\left[\frac{1}{2}\ (1/N_m + 1/N_f)\right]$，因此：

$$\frac{1}{N_e} = \frac{1}{4N_m} + \frac{1}{4N_f} \quad \text{（近似）} \qquad [4.4a]$$

$$N_e = \frac{4N_mN_f}{N_m + N_f} \quad \text{（近似）} \qquad [4.4b]$$

于是近交率为：

$$\Delta F = \frac{1}{8N_m} + \frac{1}{8N_f} \quad \text{（近似）} \qquad [4.5]$$

这提供了足够相近的近似式，除非 N_m 和 N_f 都很小，如同极度近交的情况一样。应当注意到近交率主要取决于较少的性别的数目。例如，如果每一世代中都用很大的不定数目的雌性但仅用一个雄性来保持群体，则有效数目大约只是 4。

连续世代中数目不等

同以前一样，任一世代中的近交率给定为$\frac{1}{2N}$。如果世代间数目是变化的，那么平均近交率是连续世代中$\frac{1}{2N}$的平均值。有效数目是每一世代中的数目的调和平均值。这样，在 t 世代的整个期间，

$$\frac{1}{N_e} = \frac{1}{t}\left[\frac{1}{N_1} + \frac{1}{N_2} + \frac{1}{N_3} + \cdots + \frac{1}{N_t}\right] \quad \text{（近似）} \qquad [4.6]$$

因此，最小数目的世代具有最大的影响。通过考虑等式［3.6］相应指明的新的和旧的近交，可以看出其理由。数目变大不影响以前的近交；它仅仅降低新近交的数量。因此，在一个数量波动的群体中，近交以数量变化的梯度形式而发展，而群体目前的大小只指明目前的近交率。

家系大小的非随机分布

这是偏离理想群体繁育体系的最重要离差。它的后果经常是致使有效数目低于实际的，但在特定情形下它使有效数目大于实际数目。这里家系大小是指成为下一世代繁殖个体的一个个体的后裔数目。应当记住，理想群体中每一个繁殖个体给下一世代贡献基因或后裔的概率是相等的。后裔的贡献在亲本间随机分布，家系大小因此而变化。在实际群体中，因为亲本生殖力和其后裔存活力不同，它们很少有相等的贡献后裔机会。亲本间的这种变异导致了家系大小的较大变异，并产生了这种后果，即下一世代的较大部分来自较小数目的亲本。因而减少了有效数目。相反，通过特殊的繁殖方法，家系大小的变异可能被降低至随机数量之下，但随之有效数目增加了。有效数目与家系大小变异的关系简述如下。

这里将注意力限制于具有稳定大小和雌雄个体数目相等的群体。于是所有个体，无论是雄性还是雌性个体，其平均家系大小 $\bar{k}$ 一定为 2，因为要更新群体，平均而言每一个体必须有 1 个雄性和 1 个雌性后裔作为下一世代的亲本。像理想群体中的一样，家系大小的随机变异产生了二项分布，除非 N 非常小，否则它与泊松分布几乎没有差异。泊松分布其方差等于平均数，因此当亲本对下一世代贡献机会相等时，家系大小的方差是 $V_k = \bar{k} = 2$。由于生殖力差异或其他原因，当亲本对下一世代贡献机会不相等时，则家系大小的方差大于 2。当考虑合子是同一纯合子的概率时，以类似于上一章推导近交增量的方式，可推导出家系大小的方差影响有效数目的方式。于是可从近交率中获得有效数目。这导致的关系近似为：

$$N_e = \frac{4N}{V_k + 2} \quad （近似） \qquad [4.7]$$

对 $V_k = 2$ 的理想群体，这简化成 $N_e = N$。当两性的 V_k 相同时，等式［4.7］中的关系适用于一雄一雌的交配。如果雄性能和一个以上的雌性交配，则雄性和雌性的 V_k 很可能不同。于是有效数目由下式给出：

$$N_e = \frac{8N}{V_{km} + V_{kf} + 4} \quad （近似） \qquad [4.8]$$

其中，V_{km} 和 V_{kf} 分别为雄性和雌性家系大小的方差（Hill，1979）。

大于随机数量的家系大小的变异 V_k，是 N_e 小于 N 的两个最重要的原因之一，另一个原因是随世代发生的数目变异。可采用几种不同方法来估计 N_e/N 比率。对 102 个动植物物种的 192 个估计值的综述（Frankham，1995）表明，估计值变化巨大，从太平洋牡蛎的 10^{-6} 变化到人类的 0.99。已注意到由 N_e 小于 N 的原因引起的一些变异，也获得了由总数 N 如何引起的一些变异。N_e/N 的所有估计值的平均值是 0.34，但考虑全部因素的估计值均值仅为 0.1，这比预期的低得多。甚至在物种内估计值也广泛变化，部分地取决于估计方法。例如，黑腹果蝇中，N_e/N 从 0.03 变化到 0.90，均值 0.30；在人类中，它从 0.34

变化到0.99，均值0.55。

最低程度的近交

经常希望用最低可能的近交来保持实验动物种群。尽可能多地增加繁殖个体数 N 并不是可做到的唯一事情。通过对作为亲本个体的选择，能够减少家系大小的方差 V_k，使其低于随机数量，从而增加有效数目。如果从所有家系选出的个体数相等，那么家系大小没有变异，即 $V_k = 0$。代入等式［4.7］看出有效数目近似地变为 $N_e = 2N$。精确的关系是：

$$N_e = 2N - 1 \qquad [4.9]$$

它几乎是具有相同大小的理想群体的两倍。等式［4.9］适用于由相同数目的雄雌所繁殖的群体。于是，家系大小相等意味着从每一对亲本的后裔中选择两个个体。

如果两性数目不等，通过从每一公畜后裔中选出一个雄性以及从每一母畜后裔中选出一个雌性作为亲本可以使家系大小的方差变为0。这样，近交率由下列公式给出（Gowe，Robertson 和 Latter，1959）：

$$\Delta F = \frac{3}{32N_m} + \frac{1}{32N_f} \qquad [4.10]$$

其中，N_m 和 N_f 分别为雄性和雌性亲本的实际数目，并且雌性比雄性多得多。

避免近亲如同胞或表兄妹交配，乍看似乎是一种容易降低近交率的方法。这延缓了近交的最初增加，但随后的近交率几乎没有减少。其理由先前已做过解释，等式［4.3a］给出了排除自体受精和同胞交配的有效群体大小。如果故意地使家系大小相等，那么避免近亲交配不能实现近交率的进一步减少（Robinson 和 Bray，1965）。避免近亲间交配的主要优点是使近交率在世代间较稳定，并使个体的近交系数在世代内较一致。

世代重叠

在大多数自然群体中以及家畜中，世代不是分离的而是重叠的。这意味着任何时候现存的个体都处于不同的年龄以及它们生活周期的不同阶段。而且，个体在寿命长度方面不同，继而它们繁殖的机会也不同。因此，终生寿命的差异加上生殖力的差异增加了家系大小的方差，寿命较长的个体比寿命较短的个体有更多的机会给下一代贡献后裔。等式［4.7］或［4.8］涉及了对 N_e 的影响。可是，在回答每个世代总数是多少时有一个问题，即等式［4.7］中的 N 是多少。只要群体有一个稳定的年龄结构，就可以用下列方法求出每世代总数。我们需要知道在一个特定时间间隔内出生的个体数，这个间隔可以是一年或任一适宜的一段时间。这个数目是由时间间隔定义的区组（cohort）大小。区组大小 N_c 通过 $N_c = N_T/E$ 与任何时间上的存活总数 N_T（即调查计数）相关，其中 E 是平均寿命或死亡时的平均年龄，以定义区组的特定时间间隔单位来表示（见 Emigh 和 Pollak，1979）。我们还需要知道以特定时间间隔为单位的世代长度 L，世代长度是后裔出生时双亲的平均年龄。于是每世代总数是 $N = N_c L$，每世代有效数目由下式近似给出（Hill，1979）：

$$N_e = \frac{4N_c L}{V_k + 2} \quad \text{（近似）} \qquad [4.11]$$

其中，V_k 是所有原因引起的家系大小的方差。如果雄性和雌性在数量上或世代长度

上不同，家畜中经常是这种情形，则必须以 Hill（1979）说明的方式来修改等式［4.11］。

美国人群的有效数目已被估计为 $N_e = 0.41N_T$，但比率可能在某种程度上低于这个值，因为估计值没有考虑生殖力变异的所有可能的来源（Emigh 和 Pollak，1979）。

例 4.1

来自小鼠实验的数据（Garnett 和 Falconer，1975，未发表）将用来例证本章和前一章中推导的几个公式的应用。而且，分别经方差和近交方法计算有效群体大小，我们还能检验理论的有效性。群体由 18 个系群组成，所有系群都起源于同一个随机交配的基础群体，每一世代都用 8 对亲本通过最小近交来保持所有系群（Falconer，1973）。每一系群中数据由 5 个多态酶位点上的基因和基因型频率组成。酶位点列于表 4-1。所有位点上都有两个等位基因；所有杂合体都是可辨别的并且可通过计数获得基因频率（等式［1.1］）。在第 27 世代，确定所有亲本的类型，因此在此时能无机误地确定基因频率。对每一位点都计算了 18 个系群间的基因频率方差。表 4-1 给出了每一位点的平均基因频率 $\bar{q}$、基因频率方差 σ_q^2 以及整个群体中杂合体的总频率 H。没有理由认为基因频率不同于其基础群体的值，因此对于计算都假定 $q_0 = \bar{q}$。

要进行的计算有：（1）有效群体大小 N_e，从亲本数 N 和繁殖结构中获得；（2）近交系数 F，来自基因频率方差 σ_q^2，然后计算在世代 $t = 27$ 时来自 F 的 N_e；（3）$t = 27$ 时的 F 来自杂合体频率 H，然后 N_e 再次来自 F。

1. 对于 8 对亲本，$N = 16$。当最小近交（$V_k = 0$）时，等式［4.9］给出 $N_e = 2N - 1 = 31$。等式［4.1］给出 $\Delta F = \frac{1}{2N_e} = 0.0161$，等式［3.12］给出 $F = 1 - (1 - 0.0161)^{27} = 0.355$。只有 $V_k = 0$，这些期望值才能获得。实际上一些交配对不可避免地会不育，因此 V_k 不会等于 0，从而 N_e 将低于 31。

2. F 通过等式［3.14］与 σ_q^2 相关联。以位点 *Dip - 1* 为例，$F = \sigma_q^2 / \bar{p}\,\bar{q} = 0.077/(0.236 \times 0.764) = 0.427$。每一位点给出一个独立的 F 估计值。它们在表中标题为 F（2）的栏中给出。平均值是 $F = 0.378$。根据这个 F 的平均估计值，我们从等式［3.12］可获得近交率 ΔF。经过整理，$(1 - \Delta F)^t = 1 - F_t$，$t = 27$ 时给出 $\Delta F = 0.0174$。通过等式［4.1］从 ΔF 中可以算出有效群体大小 N_e。这给出 $N_e = \frac{1}{2\Delta F} = 28.7$。

3. 在所有系群中，对相等数目的个体进行了归类，因此可从杂合体总频率 H 来估计 F。这可以通过等式［3.15］做到，但每一系群亲本数足够小，从而不能忽略期望的雄性和雌性基因频率出现差异的机会。等式［3.17］考虑到了这个因素，它给出 $(1 - F) = H/\left[2pq\left(1 + \frac{1}{2N}\right)\right]$。还以 *Dip - 1* 为例，$(1 - F) = 0.240/(2 \times 0.236 \times 0.764 \times 1.03125) = 0.646$，$F = 0.354$。而且，每一位点给出一个独立的 F 估计值，这在标题为 F（3）的栏中给出。用与上面计算 2 相同的方法从 F 中来计算 ΔF 和 N_e，这些值列入表的底部。

还可以用另一种方法（即下一章要说明的方式），从谱系记录中计算近交系数。这是一种精确的计算，因为它以实际交配产生的同一纯合体概率为基础。对每一系群都进行了计算，平均值是 $F = 0.379$。这给出 $\Delta F = 0.0175$ 和 $N_e = 28.6$。N_e（获得的）与 N_e（期望

的）的比率是 28.6/31 = 0.92，N_e（获得的）与 N 的比率是 28.6/16 = 1.79。比较表底部显示的 N_e 的 3 个估计值，我们看到来自基因频率方差的和来自杂合体频率的估计值与谱系的相当符合。可能会注意到，如果对雄性和雌性的不等基因频率来作校正，则来自杂合体频率的 N_e 的估计值是 32.7 而不是 30.3。

表 4-1 根据小鼠实验数据对若干参数的推导计算

位点	$\bar{q}$	σ_q^2	H	$F(2)$	$F(3)$
Dip-1	0.764	0.077	0.240	0.427	0.355
Id-1	0.370	0.102	0.301	0.438	0.374
Gpi-1	0.297	0.072	0.283	0.345	0.343
Gpd-1	0.215	0.042	0.253	0.249	0.273
Got-2	0.141	0.052	0.134	0.429	0.464
平均值				0.378	0.362
方法			F	ΔF	N_e
1. 根据繁殖结构预测			0.355	0.0161	31
2. 基因频率方差，σ_q^2			0.378	0.0174	28.7
3. 杂合体频率，H			0.362	0.0165	30.3
4. 谱系			0.379	0.0175	28.6

突变、迁移和选择

本章和前一章到目前为止给出的离散过程描述是以没有突变、迁移和选择等系统过程为条件的，所以它与实际群体的关联是有限的。因此，现在让我们考虑离散和系统过程联合作用时的效应。正如我们在第 2 章看到的，系统过程倾向于使基因频率以特定的值达到稳定平衡点，这些特定值对处于相同条件下的所有群体都是相同的。与之相反，离散过程倾向于使基因频率远离这些平衡值，如果这种趋势不被系统过程所阻止，则最终离散过程将导致所有基因在大小有限的所有群体中要么被固定要么被丢失。当基因频率距这个平衡值的离差愈大，系统过程使基因频率朝这个值变化的趋向就愈强烈。由于这个原因，存在一个平衡点，在这一点上系统过程阻止了基因频率的离散。于是，亚群体之间存在一定程度的分化，只要条件保持不变，这种分化就保持不变。因此，问题是要找出已达到这种平衡状态时在一个细分群体的各个系群之间的基因频率分布。但是应当注意，条件改变之后，一个群体达到平衡状态可能要花费一个相当长的时间，因此推导出的分布只能近似地适用于实际群体。解答在数学上是复杂的，将只给出主要结论。更详细的描述，见 Crow (1986)。

非频发中性突变

在 DNA 序列水平上，因为在一给定核苷酸上相同突变发生一次以上的概率很小，因此将突变考虑为非频发的是合理的；即，在每一个位点上都有可潜在发生的数目无穷大的有效突变。这是突变的无穷等位基因模型，其结果是一种突变在一个群体中只出现一次。

首先我们将简要考察罕见突变的最终结果，这种突变对于适合度来说是有效的，中性

的。这类突变体的最终结果形成了分子进化的中性理论基础，Kimura（1983）对此给予了一个全面的考察。常染色体位点上的每一个突变体只有一半的机会在下一代中存活，并且大部分突变体在几个世代之后就消失了，这一点在第 2 章已有解释。可是，少数突变体将存活下来，并且通过离散过程在群体中分散开来。其中一些突变体后来消失了，但有一些将最终被固定。在早期世代中由突变产生的一个新等位基因的固定被称为“等位替代”。在一特定位点上发生等位替代的机会是多少？在 N 个个体的一个二倍体群体中，有 $2N$ 个等位替代位点的代表（representatives），当没有选择时，它们当中的一个最终被固定。因此，任一特定基因被固定的机会是$\frac{1}{2N}$。用 u 表示所讨论位点的中性突变率，即在任一世代中出现由突变产生的一个新的中性等位基因的概率。假定每一个新的突变体开始只有一个拷贝，那么这个位点上新突变体的总数为 $2Nu$。单独对每一个突变体，固定的机会是$\frac{1}{2N}$。因此，一个或另一个新突变体被固定的概率是 $2Nu \times \frac{1}{2N} = u$。这就是在任一特定世代在一个位点上出现的等位替代概率，它简单地等于所讨论位点上每个世代的突变率。为了一起获得所有位点上的替代率，我们使 u 等于每一配子的中性突变率，即任一位点上携带一个新的中性突变体的配子频率。为什么等位替代率与群体大小无关，其理由是，在一个较大的群体中较大数目的新突变体被较小的个体存活机会所平衡了。

当然，根据新突变体是有利的还是不利的，选择可增加或减少固定的机会。预计相当多的突变体是有害的而不是有益的。一个有害突变体引起等位替代的机会是多大？通过同样的推导，这等于“有效中性”等位基因的突变率；按照 Kimura（1983）的描述，一个“有效中性”等位基因是在 $s = 0$（即严格中性）至 $s = \frac{1}{2N_e}$范围内取值，具有不利于自身的选择系数 s 的一个等位基因。因此，有效中性取决于有效群体大小，对一个有效中性等位基因来说，乘积 $N_e s$ 小于$\frac{1}{2}$。

预计一些有害突变体因随机漂变而被固定，这意味着经过长时间所有群体必须倾向于降低自然适合度，除非决定适合度的条件改变了或者下降被有利突变体的出现和固定所平衡了（见 Kimura，1983，p.248）。这个预测具有重要的实际结果。实验动物的近交系可通过兄妹交配来保持，其有效群体大小大约为 2.5（等式 [4.2a]）。因此有害突变体其选择系数达到大约 20% 时就会被固定，随之而来繁殖适合度的损失是巨大的，除非通过强烈的人工选择抵消它；有利突变体极其稀少以致不能抵消小群体中适合度的损失。Wright 和 Kerr（1954）用果蝇（*Drosophila*）的一个实验例示了小群体中一个有害基因的固定。

频发突变和迁移

频发突变和迁移可以放在一起处理，因为它们以相同的方式改变基因频率。再次考察一个细分为很多系群的群体，所有系群的有效大小为 N_e；用每一系群每一世代的部分繁殖个体 m，表示随机来自所有其他系群的迁入者。用 u 和 v 表示一个位点上两个等位基因之间两个方向的突变率。当一方面的离散与另一方面的迁移和突变之间达到平衡时，亚群体之间离散的状态可以如下的近交系数表示：

$$F = \frac{1}{4N_e(u + v + m) + 1} \quad （近似） \qquad [4.12]$$

如果平均基因频率已知，那么通过代入等式［3.14］的 $\sigma_q^2 = F\bar{p}\bar{q}$，可以按基因频率方差表示离散状态。

图 4-1 显示了对应于 F 的 4 个平衡值的基因频率理论分布。这些分布在一般形式上类似于图 3-4 显示的、没有突变和迁移的群体在近交过程期间所经历的分布。突变或迁移的效应被认为是在某一值点上离散过程中止，该值点相应于 F 或亚群体间基因频率方差 σ_q^2。这里主要差别是，如果 F 超过 0.33，当所有基因频率包括固定都是同样地可能时，分布变成了 U 形，更多的亚群体处于极端基因频率，较少的亚群体处于中间基因频率。其原因是，在发生突变或迁移时任一系群中的固定都不是永久的。

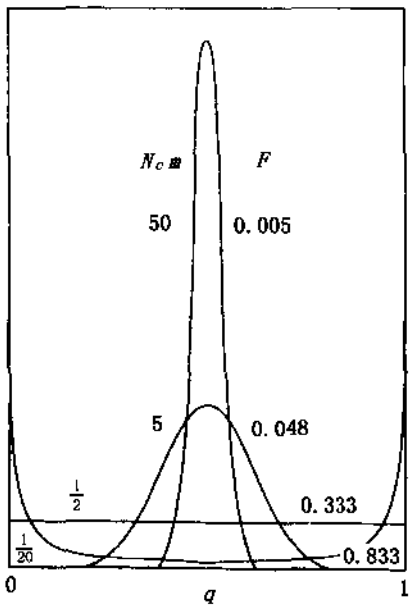


图 4-1 当离散过程与突变或迁移相平衡，且平均基因频率为 0.5 时，亚群体之间基因频率的理论分布

垂直轴是概率，与图 3-4 的相同。曲线指示的离散状态由图中的 F 值指明。$N_e m$ 的值是每个世代迁移者的数目，正像正文中解释的那样（根据 Wright，1951）。

为了明白等式［4.12］和图 4-1 的分布意义，让我们考虑这些问题：如果突变率已知但没有迁移，那么群体在怎样的 F 值上稳定？没有突变，需要多少迁移才能产生所显示的分布？为此目的，把等式［4.12］作为 $N_e = \dfrac{1}{4F(u+v+m)}$ 将是足够准确的。用通常的突变率 $u+v=10^{-5}$、$m=0$ 以及对应于图 4-1 最小离散分布的 $F=0.005$ 代入该表达式。这给出 $N_e = 5\times10^6$，这意味着在被突变稳定以前，在 $F=0.005$ 的范围内，将分化为大小为 5 百万的亚群体。较小的亚群体将进一步分化。例如，对应于 $F=0.333$ 的均匀分布将分化达到大小为 75 000 的亚群体。对很多物种，可能是大多数，就亚群体内随机交配的亚群体来说，甚至这是一个不合实际的巨大规模。因此，结论是频发突变作为延缓或阻止由随机漂变引起的亚群体分化的一个因子是可以忽略不计的。可是，迁移则很不相同。用 $u+v=0$ 重新整理等式［4.12］，给出 $F=\dfrac{1}{4N_e m+1}$。因此，离散状态取决于乘积 $N_e m$，它是每个世代迁入者的实际数目，而与群体大小无关。这个结论开始可能显得似乎不合理，若注意到较小群体比较大群体需要较高的迁入率来保持相同的离散状态便可以理解这个结论了。代入对应于图 4-1 中分布的 F 值就得出进入图中的 $N_e m$ 的值。例如，每两个世代一个迁入者的平均值 $\left(N_e m=\dfrac{1}{2}\right)$ 就足以保持对应于 $F=0.333$ 的基因频率的平坦分布。结论是很小数目的迁入者就可阻止由随机漂变引起的很大分化。为什么突变和迁移的效应差异如此大的原因与第 2 章指出的相同：实际突变率远远小于实际迁移率。

前面考虑的迁移所涉及的情形被称为“岛模型”。它描绘了一个不连续的群体，任一亚群的迁入者都有相等的概率来自任何其他亚群。一个更实际的模型是“邻域模型”或“距离隔离”。群体被描绘为连续分布在居住的地域上，但被个体在出生和繁殖之间移动的

有限距离细分为“邻域”。邻域就是一个区域，在该区域内交配是有效随机的，这相当于一个亚群体。可是，基因频率在地域之间的邻域间连续变化。因为邻域的迁入者来自邻近的通常比来自较远的要多，所以迁入者基因频率的差异要小于岛模型中迁入者的。因此迁移在抵消随机漂变方面就不那么有效。邻域模型导致的结论是：如果邻域中有效数目数量级为20，则将产生巨大数量的局部分化，如果数量级为200，则产生中等数量的局部分化，如果数量级大于1 000左右，所产生的局部分化则可忽略不计。

选 择

对一个大群体内的一个位点施加选择，当选择有利于杂合体时，可使基因频率以中间值达到平衡，当选择被突变平衡时，可使基因频率以低值达到平衡。离散过程倾向于使基因频率偏离其平衡值。这降低了群体的平均适合度，因为处于平衡时负荷最小，而且一些亚群体的有害等位基因甚至可能被固定。基因频率偏离平衡值越远，选择的效果就越强。因此，选择和随机漂变的对抗作用达到平衡时，在平衡点处亚群体间基因频率有一个稳定的分布。这时问题是：借助它们与最佳基因频率的随后离差，亚群体必须多小才能产生可察觉的分化？下面说明的情形必将足够回答这个问题，但要理解突变、选择和离散的联合效应，读者必须查阅其他原始资料。

首先考察有利于杂合体的选择。效果取决于平衡的基因频率。当两个纯合体处于相等的劣势，因而平衡的基因频率为0.5时，基因频率的分布看起来大致像图4-1中的那样。在 $N_e = 1\,000$的亚群体中，通过不利于两个纯合体的选择系数 $s = 0.1$ 可以达到相应于 $F = 0.005$ 的最小离散的分布。更大的离散将需要较少的选择或较小的群体。具有相当数量固定的最大离散分布将大致需要 $s = 0.01$ 和 $N_e = 100$。因此，对于杂合体的选择不允许大量的随机漂变，除非选择很弱（大约1%）或群体规模很小（大约100）。可是，如果平衡基因频率不是0.5，选择在阻止随机漂变中就不那么有效。当平衡基因频率大约高于0.8或低于0.2时，选择实际上加速了随机漂变（Robertson，1962）。其原因是那时一个纯合体比另一个的适合度大得多，选择增加了适合度较强的纯合体被固定的概率。

下面考察不利于一个隐性等位基因的选择因频发突变而平衡的情形。这是很难说明的，因为对于选择系数的实际值，平衡基因频率会非常低，分布被挤拢到接近 $q = 0$ 的极限情形。图4-2显示了平衡基因频率大约为 $q = 0.2$ 时很弱选择下的3种稳定分布，认为两个方向上的突变率相同，选择系数 s 是突变率的20倍。如果我们假定突变率为 10^{-5}，那么 $s = 20 \times 10^{-5}$，分布涉及的群体大小是（a）250 000、（b）25 000和（c）2 500。同样结论还是，选择不允许大量的随机漂变，除非选择很弱或群体规模很小。随机漂变的大小大约取决于群体大小和选择系数的乘积 $N_e s$，图4-2中3种分布对应的 $N_e s$ 值分别等于50、5和0.5。

自然群体中的随机漂变

已描述了离散过程和它的理论结果，现在我们可以转向更实际的问题，即这些结果在自然群体中实际能见到多少。这个问题的回答有重重困难，下面的评论更多地是用来表明这些困难的本性，而不是去直接回答这个问题。

本章和前一章简述的小群体理论，本质上主要是数学的理论，毫无疑问是有效的：只

要给出孟德尔遗传的机制，获得的结论便是指定条件下的一种必然结果。于是所讨论的问题是自然群体中的条件是否经常允许基因频率的离散成为可查明的。如果条件适宜，由离散过程预计的现象是不同地区居住者之间的分化以及连续世代之间的差异。这两种现象在细分或小的隔离群体中都是众所周知的，引出的结论是，因为它们是随机漂变的期望结果，所以随机漂变一定是它们的原因。但还有其他可能的原因：在不同地区以及不同季节环境条件可能不同；因此，选择强度，甚至选择方向也因地因年而变化很大，并且观察到的差异同等程度上可归因于选择压的变异。因此，在我们能够合理地将这些现象归因于随机漂变之前，我们必须了解：（1）有效群体大小足够小；（2）亚群体被充分地隔离（或“邻域”的大小足够小）；（3）涉及的基因极少受选择的影响。

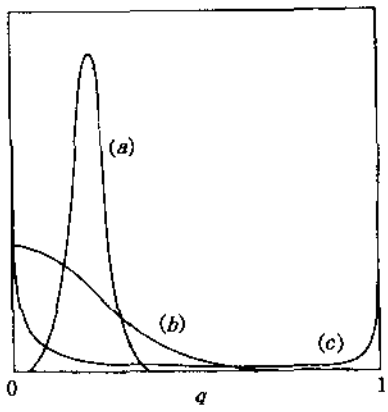


图 4-2　当离散过程因突变和选择而平衡时，亚群体之间基因频率的理论分布

图显示出在群体大小为：$(a)N_e = 50/s$，$(b)N_e = 5/s$ 和 $(c)N_e = 0.5/s$ 中的 $u = v = \frac{1}{20}s$ 的一个隐性基因（根据 Wright，1942）。

对群体现有大小的估计，虽然技术上不容易，原理上却没有困难。但现有的分化状态取决于过去的群体大小，而这一般只能猜测。群体可能经历了一个或多个瓶颈——在瓶颈世代中，它的大小被不利条件剧烈减少。由很低数目的这些世代引起的离散是永久的，并不被群体大小的随后增加而消除。（等式［4.6］表明有效群体大小如何受变化数目的影响。）如果一个物种开拓了一片新的领域，新的亚群体的建立成员可能在数量上很少，却引起了第一个世代中巨大数量的随机漂变。这被称为*创造者效应*。如果亚群体接着扩大，则它与主要群体的差异可能显得太大以致于不与它现有的数目相一致。把这些差异归因于创造者效应经常是合理的，但在谱系记录缺乏的情况下，它也只能是一种猜测。作为创造者效应的一种结果，隔离群体正在分化的最清晰和最有趣的例子之一，是在美国 Amish 社区见到的，McKusick（1978）对其进行了研究，通过谱系记录建立了创造者效应。

判断亚群体是否被充分隔离并不很难。对一个不连续的群体，我们在很大程度上有理由确信没有太多的迁入者；对一连续群体，“邻域”的大小至少在原理上是可以度量的。最大的困难在于估计对有关基因起作用的自然选择的强度。强度远低于实验可查明的选择足以阻止除最小群体外的所有群体中的离散。与酶多态性有关的很多基因可能有足够低的选择系数以致于允许群体分化（见下面）。与数量差异有关的基因对要发生的随机漂变也极可能是中性的。毫无疑问，所有这类基因在实验群体中确实表现了随机漂变，这一点将在以后各章说明。

多态性

在第 2 章我们看到，通过选择力量的平衡，有几种途径能够保持多态性，但这些似乎不象是所有多态性的理由。一种可替换的选择模型是，在选择不起作用的情况下突变和随机漂变之间的一种平衡。现在将描述这种模型。

中性理论

分子进化的中性理论假设大多数新的突变是有害的，只有很少是有利的，并且一些部分是选择中性的（即、对适合度没有影响）。有害的和有利的突变将分别被很快地消除和固定，对多态性几乎没有贡献。在任何时候，群体中的变异将主要由中性等位基因引起，这些基因注定要最终被固定或消失（Kimura，1983）。因此，这种模型下的多态性被认为是短暂的。但是，中性等位基因能在群体中持续很长时间。对一个注定要被最后固定的中性等位基因来说，平均而言，这是 $4N_e$ 个世代。注定要消失的中性等位基因不会持续像平均长度那么长的时间，但消失的时间有很高的方差。

经受新的独一无二的中性突变持续输入并且因漂变令现有变种固定或消失的一个群体，最终将达到一种稳定状态，此刻变异消失的速度精确地为增加的速度平衡。虽然等位基因精确的同一性将随等位基因来自和离开群体的全部时间而变化，但处于突变—漂变平衡的群体将具有等位基因频率的分布以及作为有效群体大小和中性突变速率函数的平均杂合性。通过修改等式［3.6］以包括突变，来自中性突变和随机漂变平衡的平均杂合性是易于获得的。根据突变的无限等位基因模型，在大小为 N_e 的一个群体中，每一个等位基因以速率 u 只产生一次。因此任何等位基因的纯合体是血统同一的。在一个有限群体中，携带同一的中性等位基因的两个配子的概率由等式［3.6］给出。只要两个等位基因都不突变，则这两个等位基因仍然保持同一，其概率为$(1-u)^2$,因此：

$$F_t = \left[\frac{1}{2N_e} + \left(1 - \frac{1}{2N_e}\right)F_{t-1}\right](1-u)^2 \qquad [4.13]$$

处于随机漂变和突变之间平衡，$F_t = F_{t-1}$。为 F_t 和 F_{t-1} 代入平衡值 $\hat{F}$，重新整理得到：

$$\hat{F} = \frac{(1-u)^2}{2N_e - (2N_e - 1)(1-u)^2} \qquad [4.14]$$

每一个位点的突变率 u 很小，因此，u^2 和 $2u$ 项能被忽略。于是等式［4.14］可被简化成：

$$\hat{F} = \frac{1}{4N_e u + 1} \quad （近似） \qquad [4.15]$$

$\hat{F}$是一个群体中由突变—漂变平衡引起的一个中性等位基因的期望纯合性。平衡的杂合性 $\hat{H}$是 $1-\hat{F}$：

$$\hat{H} = \frac{4N_e u}{4N_e u + 1} \quad （近似） \qquad [4.16]$$

这也是对所有位点的平均期望杂合性。在第 2 章，一个多态性位点被定义为最共同的等位基因的频率不超过 0.99 的位点。于是，共同等位基因和任何其他的等位基因的杂合体的频率是 0.02，这是 H，因为两个稀有等位基因的杂合体频率是可以忽略的。这是要解释的 H 的最低值，因为具有较低值的位点不会被列为多态性的。对 $H=0.02$，等式［4.16］表明 $4N_e u$ 也近似地为 0.02。因此，在这种模型下，如果中性突变率的倒数大于有效群体大小的 5 000 倍，多态性将被保持。

检验　中性理论的有效性已被很多研究检验，关于它们，这里仅给出简要的概括。给

定 N_e 和 u 值，并假定无限等位基因模型，中性理论导致了如下可观察的多态性特性期望值：杂合性、多态性位点的比例、群体内位点间等位基因频率的分布、群体间或亲缘紧密物种之间的一个位点上等位基因频率的分布（Kimura，1983）。通过比较这些参数的观察值和期望值，可对中性理论进行评估。

让我们首先考察来自观察的杂合性的推断结果。对很多位点、群体和物种，存在关于等位酶频率的非常庞大的数据，可以用它们来检验与中性期望值的符合程度。例如，对所有研究的物种，等位酶杂合性从 0 变化到 0.30，具有近似为 0.05 的众数（Kimura，1983，p.254）。考虑到 10 次突变中只有 1 次是中性的，剩下的都是有害的，Kimura（1983）证明等位酶的中性突变率可能是 10^{-7}的一个数量级。对 N_e 为 10^5，于是 H 近似地为 4%，这接近于观察到的众数。问题是要知道 N_e 的真值是多少，因为对 N_e 和 u 的一个适宜选择，在中性理论下可以调节 H 的任何值。等位酶位点的中性突变率似乎在物种之间没变化。因此，不同群体的杂合性应当随着它们的群体大小而变化。生物体在群体大小方面的变化可达到几个数量级的大小，但在被研究的物种中，杂合性的不同很少超过 10 这样一个因子。这种差异是表明等位酶变种受制于选择的一个原因（Lewontin，1974）。可是，从种群调查中获得的群体大小不是 N_e 的准确估计值，正如本章先前提到的那样。在过去，特别是在物种形成的时候，作为瓶颈的结果，物种之间 N_e 的变异比它出现时要小得多（Kimura，1983）。

除经常在自然群体中发现的一些额外的稀有等位基因之外，观察到的等位基因频率分布一般与中性期望值吻合得很好。中性理论的其他统计学检验，是以等位基因的数目和杂合性之间、杂合性的平均值和方差之间，以及多态性位点的比例和杂合性之间的期望关系为基础的。对等位酶数据，与中性理论有着良好的但不是完全的吻合。可是，注意到检验是以严格中性为基础的，而样本含有包括低频率的有害等位基因在内的所有等位基因。因此与严格中性的有某种偏离是可以预料的，与中性期望值比较，观察到的额外的稀有等位基因与中性理论不是不一致的。对来自中性多态性的选择的识别，更强有力的检验是比较 DNA 序列。中性理论预言，与氨基酸替换多态性相比，沉默位点（同义密码子）和蛋白质非编码区的序列多态性应当更频繁，因为对于沉默位点和内含子来说，所有是中性突变的片段要比由氨基酸替换引起的突变片段多得多。这是“选择强制”的概念，观察和这些期望是一致的。中性理论还预言物种内的杂合性与物种间的趋异性之间存在相关，这与来自1 500个物种的数据是一致的（Skibinski，Woodwark 和 Ward，1993）。这些检验以及结果览要的更详细描述参见 Kimura（1983）以及 Hartl 和 Clark（1989）。

在多年激烈地争论之后，结论必然是很多分子多态性归因于中性突变和漂变的平衡，虽然一些多态性无疑是适应性的且被选择力量的平衡所保持。因此没有单一机制负责保持遗传变异；确切地说，是不同的力量按照它们的功能以及在遗传组织的不同水平上对位点产生作用。因为多态性位点负责数量性状的变异，因此，我们将在第 20 章关于数量性状变异性保持的段落中回过头来讨论分子多态性的近中性以及表型的自然选择问题。

习 题

4.1 假定通过在一个瓶子中放一固定数目的未交配成年果蝇（*Drosophila*）并允许它们随机交配来保持 4 个果蝇种群。所有种群有 10 个雌亲，但雄亲数不同，雄亲数分别为

10、5、2 和 1。计算每一种群的有效群体大小和 10 个世代之后的近交系数。假定雌亲之间以及雄亲之间没有生殖力的差异。

[解答序号：83]

4.2 繁殖个体的性比率可以表示为每一雄性拥有的雌性数。修改等式［4.4*b*］以便能够用雌性数 N_f 和每一雄性的雌性数 d 的形式来表示 N_e。

[解答序号：93]

4.3 假定一个隔离的自然群体经历一个有规律的 5 年循环，在连续世代中参加配种的对数为 500、50、100、200 和 400。有效群体大小和近交率是多少？

[解答序号：103]

4.4 当一种植物的两个变种，一个是自花可育的，另一个是自花不育的，在 20 株单个植物之间通过随机授粉繁殖时，比较这两个变种的（近似）近交率。

[解答序号：113]

4.5 计划用每世代 8 对小鼠交配并用最小的近交来保持一个小鼠种群。可是，这个计划不能严格坚持，因为一些交配对未能提供所需要的 2 个后裔。在一个特定的世代中，8 对交配提供的作为亲本的后裔数如下：0、1、1、2、2、3、3、4。该世代中有效群体大小是多少？

[解答序号：123]

4.6 习题 3.6 中描述的小鼠种群每一系群的育种计划是使 8 对个体交配并使用来自每一对的 2 个后裔作为下一世代的亲本。如果严格坚持这项计划，那么系群的有效群体大小是多少？习题 3.6 中数据显示的有效群体大小实际是多大？

[解答序号：133]

第 5 章　小群体：Ⅲ．系谱群体和近亲近交

在前面两章中，通过繁殖个体的有效数目描述了小群体的遗传特性；并且推导出了用有效数目形式表示的表达式，借助这些公式，基因频率的离散状态可表示为近交系数。近交系数，它系任一个体是同一纯合体的概率，可从群体大小和特定的繁殖结构中推导出来。因此，它表示了一个世代中所有个体的平均近交系数。可是，当个体的系谱已知时，近交系数可以更方便地直接从系谱中推导出来，而不必间接地从群体大小中去推导。这种方法在实践中有几个优点。经常需要知道的是个体的近交系数，而不是整个世代的近交系数，以及依系谱计算所产生的结果是什么。在家畜中，一些个体经常作为亲本出现在两个或更多的世代中，当系谱已知时，世代的这种重叠不会造成什么麻烦。因此，本章中考虑的第一个题目是从系谱中计算近交系数。第二个题目涉及近亲近交的规则系统。当自体受精被排除时，用群体大小形式表达的近交率只是一个近似值，而如果群体大小很小，则近似式并不够近似。因此，在近亲近交的系统下，近交率必须用不同的方法去推导．借助系谱也能最好地做到这一点。

当近交系数从真实群体的系谱中推导时，它不必描述基因频率的离散状态。它基本上是关于系谱亲缘关系的一种陈述，而它与离散状态的对应取决于抵消离散过程的缺乏，特别是取决于可以忽略的选择。在前面几章中，我们能够使用近交系数作为离散过程的度量，因为它与基因频率方差的关系的必要条件已被指定。

系谱群体

一个个体的近交系数

这个系数是产生于血统同一个体的配子所携带的一对等位基因的概率。因此近交系数的计算仅仅需要将系谱追溯到双亲的共同祖先，且计算每一次分离的概率。考虑图 5-1 中的简单系谱，它代表半同胞之间的交配。X 是我们想知道其近交系数 F_X 的个体。它的双亲 P 和 Q 通过它们的共同亲本 A 而有亲缘。它们在任何其他方面都无亲缘，因此我们只需考虑 A 的基因通过 P 和 Q 传递到 X，并计算 X 是同一纯合体的概率。假定用符号 A_1 和 A_2 表示 A 在任一特定位点上所携带的基因。X 是 A_1A_1 的概率是$\frac{1}{16}=\left(\frac{1}{2}\right)^4$，因为 A_1 通过 4 条通径 AP、PX、AQ、QX 的每一条被传递下去的机会对每一条通径都是 1/2。类似地，X 是 A_2A_2 的概率是$\left(\frac{1}{2}\right)^4$，X 是 A_1A_1 或 A_2A_2 的概率是 $2\left(\frac{1}{2}\right)^4=\left(\frac{1}{2}\right)^3=\frac{1}{8}$。X 是同一纯合体的这种概率代表由作为 P 和 Q 共同祖先的 A 所产生的新的近交。可是，通过以前的近交，共同祖先 A 本身可能就是一个同一纯合体，在这种情况下，X 也将是一个同一纯合体，如果它获得基因型 A_1A_2 或 A_2A_1（根据 A_1 是来自 P 还是来自 Q 可区别这两种基因

型)。依据与以前相同的原因，这些基因型每一个的概率都是$\left(\frac{1}{2}\right)^4$，并且一个或另一个的概率都是$\left(\frac{1}{2}\right)^3$。A 是同一纯合体的概率就是它的近交系数 F_A。于是，由于以前的近交，X 是同一纯合体的附加概率是$\left(\frac{1}{2}\right)^3 F_A$。将两个部分的近交放在一起给出 X 的近交系数为 $F_X=\left(\frac{1}{2}\right)^3+\left(\frac{1}{2}\right)^3 F_A=\left(\frac{1}{2}\right)^3(1+F_A)$。注意指数 3 是通过它们的共同祖先将亲本联系起来的通径中的个体数，即个体 P、A 和 Q。这使得通过简单地计数通径中的个体可容易地求解出概率。在图 5-1 中，只有双亲和共同祖先；在图 5-2 中，共同祖先进一步向上追溯，要计数的个体是 P、D、B、A、C、Q，组成 6。因此，图 5-2 中的 F_X 是$\left(\frac{1}{2}\right)^6(1+F_A)$。在较复杂的系谱中，双亲可以由于一个以上的共同祖先，或通过不同通径的同一个共同祖先而彼此有亲缘，这一点示例在例 5.1 中。于是，每一个共同祖先和每一个通径都对后代是同一纯合体贡献一个附加概率，将使双亲有亲缘的每一个通径的各个概率相加起来可获得近交系数。

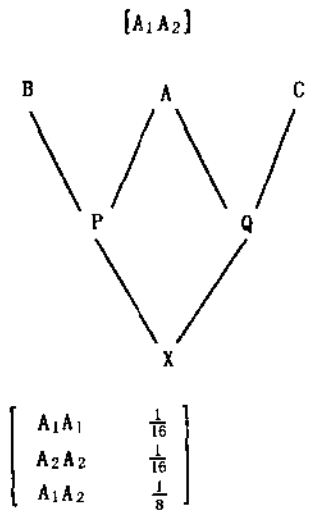


图 5-1

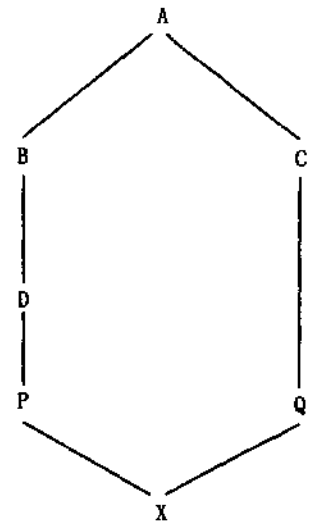


图 5-2

将所有这些加在一起就给出如下的一个个体近交系数的一般公式：

$$F_X=\sum\left(\frac{1}{2}\right)^n(1+F_A) \qquad [5.1]$$

其中 n 是在任一亲缘关系通径中的个体数目，计数 X 的双亲、共同祖先以及连接双亲和共同祖先的通径中的所有个体；对所有亲缘关系的通径求和。当采用这种方式计算近交系数时，定义被提及的现有近交的基础群体是必要的。基础群体中个体的近交系数被指定为 0。实际上，可以简单地认为基础群体的个体就是位于系谱顶部的那些个体，进一步向上追溯其祖先是不知道的。

例5.1

图5-3中的系谱将说明公式［5.1］的使用。要计算近交系数的个体是X。我们必须找出使X的双亲P和Q彼此有亲缘的通径。对这种亲缘关系没有贡献的通径用虚线表示。假定除了显示的这些个体以外，任何个体之间都不存在亲缘关系。有4个个体是共同祖先，即A、B、F和J，造成P和Q之间的亲缘关系。亲缘关系的通径以及 F_X 的计算（四舍五入到4位小数）显示在表5-1中。求出的X的近交系数0.0606。下面几点应当注意：（1）D和E是全同胞。它们的亲缘关系造成了X的亲本之一P的某种程度的近交，但它没有造成P和Q之间的亲缘关系，因此对 F_X 没有贡献。（2）E和F是半同胞，因此作为P和Q共同祖先之一的J的近交系数，与图5-1中的一样，为1/8。（3）有4条通过B作为共同祖先将P和Q联系起来的通径，所有这4条通径都必须包括在计算中。（4）在同一条通径中没有一个个体能出现两次。例如，P M J E **B** F J N Q不是一条有效的通径，因为它产生的近交完全被计算在较短的通径P M J N Q中J的近交系数之内。（5）最后，必须小心不要以错误的方向经过通径：例如，F不能通过K和N将基因传递到P。

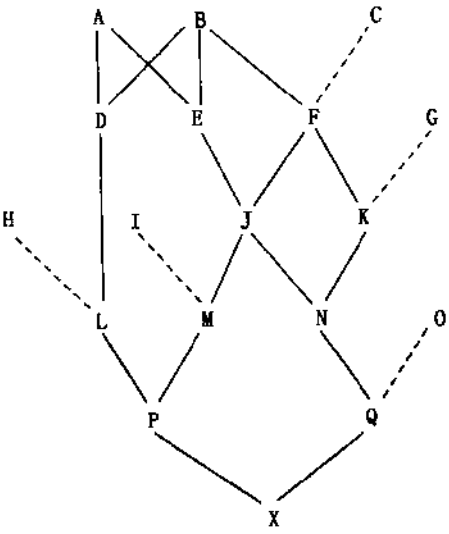


图5-3

表5-1 亲缘关系通径以及 F_x 的计算

亲缘关系通径	n	共同祖先的 F	对 F_X 的贡献
P L D A E J N Q	8	0	$\left(\frac{1}{2}\right)^8 = 0.0039$
P L D B E J N Q	8	0	$\left(\frac{1}{2}\right)^8 = 0.0039$
P L D B F J N Q	8	0	$\left(\frac{1}{2}\right)^8 = 0.0039$
P L D B F K N Q	8	0	$\left(\frac{1}{2}\right)^8 = 0.0039$
P M J E B F K N Q	9	0	$\left(\frac{1}{2}\right)^9 = 0.0020$
P M J F K N Q	7	0	$\left(\frac{1}{2}\right)^7 = 0.0078$
P M J N Q	5	$\frac{1}{8}$	$\left(\frac{1}{2}\right)^5 \times \frac{9}{8} = 0.0352$
			$F_X = 0.0606$

当系谱长并复杂时，追踪亲缘关系的所有通径可能是不切实际的。但是，通过抽取有限数目的通径能够得到足够准确的近交系数估计值（Wright和McPhee，1925）。

共祖率或亲缘关系

还有另一种计算近交系数的方法，它经常是更方便的，并且更容易适应各种问题。在下一部分它将用于求解规则的近亲近交系统下的近交系数。它在实践中的主要用途是设计产生最小近交的交配以及计算一个完全的系谱群体中逐个世代的近交系数。这个方法在原理上与上面给出的公式［5.1］没有什么不同，但不是从目前的个体追溯到共同祖先，而

是一直向前追溯，逐个世代地记录系谱，并计算出由现在进行的交配将产生的近交。一个个体的近交系数取决于两个亲本方面的共同祖先的数目。因此，我们不考察后代的近交，我们可以去考察两个亲本之间由血统造成的亲缘关系程度。这被称为*共祖率*或*共祖系数*或*血缘系数*，将用符号 f 来表示它。如果任何两个个体交配，则这两个个体的共祖率与它们后代的近交系数完全相同。因此，两个个体的共祖率是两个配子携带同一血统等位基因的概率，这两个配子是随机获取的，并且每个个体提供一个。

考察图 5-4 中的一般化系谱。X 是亲本 P 和 Q 以及祖亲 A、B、C 和 D 的一个个体。现在，P 和 Q 的共祖率完全由亲缘 A 和 B 与 C 和 D 的共祖率来确定，并且如果这些共祖率是已知的，我们就无需进一步追溯系谱。可以看出 P 和 Q 的共祖率简单地就是 AC、AD、BC 和 BD 4 个共祖率的平均值。如果采用概率的形式来表示，这将更明显，虽然用言语来表示时解释是繁琐的。随机地从 P 和 Q 中各取出一个配子并重复多次。在一半的情形中，P 的配子将携带 A 的一个基因，在另一半，将携带 B 的一个基因；Q 的配子也是类似的情况。因此，两个配子，一个来自 P，一个来自 Q，有四分之一的情形将携带来自 A 和 C 的基因，有四分之一的情形将携带来自 A 和 D 的基因，有四分之一的情形将携带来自 B 和 C 的基因，有四分之一的情形将携带来自 B 和 D 的基因。现在，随机取自 A 和 C 的同一血统的两个配子的概率就是 A 和 C 的共祖率，即 f_{AC}，其余类推。因此，现在恢复成符号，

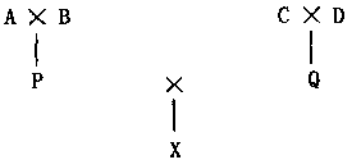


图 5-4

$$f_{PQ} = \frac{1}{4}f_{AC} + \frac{1}{4}f_{AD} + \frac{1}{4}f_{BC} + \frac{1}{4}f_{BD}$$

这给出了使一个世代的共祖率与下一世代的共祖率有亲缘的基本规则：

$$F_X = f_{PQ} = \frac{1}{4}(f_{AC} + f_{AD} + f_{BC} + f_{BD}) \qquad [5.2]$$

借助这个规则，实验者可将逐个世代的共祖率制成表，这为计划交配和计算近交系数提供了基础。Plum（1954）给出了更详细的操作解释。

如果存在世代重叠，我们可能需要找出属于不同世代的个体共祖率，为此需要一个补充规则。对概率的考察表明两个个体的共祖率等于一个个体和另一个个体的两个亲本的平均共祖率。因此，参考同一系谱（图 5-4），给出 P 和 C、P 和 D 共祖率的规则是：

$$\left.\begin{aligned} f_{PC} &= \frac{1}{2}(f_{AC} + f_{BC}) \\ f_{PD} &= \frac{1}{2}(f_{AD} + f_{BD}) \end{aligned}\right\} \qquad [5.3]$$

这个规则也给出：

$$f_{PQ} = \frac{1}{2}(f_{PC} + f_{PD})$$

通过等式［5.3］的代入，就简化成等式［5.2］的基本规则。

在我们能将这种方法应用于系谱群体或规则的近交系统之前，我们需要知道一些共祖率的数值。第一代的双亲必须被假定为全无亲缘，$f = 0$。它们的后代之间有最初的非零共祖率，并且当所有这些共祖率都被确定，所有随后的世代便可根据上面给出的规则来计算了。第一世代中可能需要共祖率的亲缘关系为亲子、全同胞、半同胞和自己。在下一部分

为求解连续近交的结果也需要这些亲缘关系的共祖率。因为这些出现在所有其他的情形中，从自己开始的共祖率如下。

自　己　一个个体与它本身的共祖率 f_{AA} 是通过自交产生的后代的近交系数。这是随机从 A 获取的两个携带同一等位基因的配子的概率，由于下面的原因，它是 $\frac{1}{2}(1+F_A)$。假定 A 的基因是 A_1 和 A_2。随机获取的两个配子都是 A_1 或都是 A_2 的概率是 1/2。一个配子是 A_1，另一个配子是 A_2 的概率是 1/2，但是，A_1 和 A_2 是同一血统的概率是 A 的近交系数 F_A。这样，两个携带同一等位基因的配子总概率是 $\frac{1}{2}+\frac{1}{2}F_A$，因此：

$$f_{AA}=\frac{1}{2}(1+F_A) \qquad [5.4]$$

如果已知（或假定）F_A 为 0，那么 $f_{AA}=\frac{1}{2}$。

亲子　它们处在不同的世代中，因此补充规则［5.3］是适用的。在图 5-4 中，P 和 A 的共祖率等于 P 的双亲和 A 的平均共祖率，即：

$$f_{PA}=\frac{1}{2}(f_{AB}+f_{AA}) \qquad [5.5]$$

如果已知或假定 A 和 B 无亲缘以及 A 不是近交的，那么 $f_{AB}=0$，$f_{AA}=\frac{1}{2}$，共祖率简化成 $f_{PA}=\frac{1}{4}$。

全同胞　它们处在同一个世代中，因此基本规则［5.2］适用。如果系谱写成图 5-5 那样的形式，那么这个规则的应用更容易理解。A 和 B 既是 P 又是 Q 的双亲，P 和 Q 是全同胞，具有一个后代 X。应用基本规则［5.2］并注意到 $f_{AB}=f_{BA}$，有：

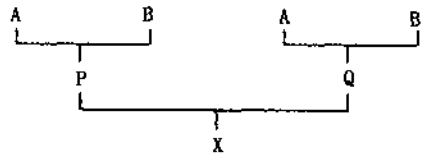


图 5-5

$$f_{PQ}=\frac{1}{4}(2f_{AB}+f_{AA}+f_{BB}) \qquad [5.6]$$

由于以前没有近交或亲缘关系，简化成 $f_{PQ}=\frac{1}{4}$。

半同胞　图 5-1 给出了半同胞的一个系谱。应用基本规则［5.2］，并注意 A 同时是 P 和 Q 的一个亲本，给出：

$$f_{PQ}=\frac{1}{4}(f_{AB}+f_{AC}+f_{BC}+f_{AA}) \qquad [5.7]$$

由于以前没有近交或亲缘关系，简化成 $f_{PQ}=\frac{1}{8}$。这个结果在图 5-1 中已经作为 P 和 Q 的后代 X 的近交系数而得到了。

规则的近交系统

一个规则的近交系统是这样一个系统：其中同一交配制度可在所有世代中应用，而且同一世代中的所有个体具有相同的近交系数。规则的系统最经常用于产生快速近交，因此交配是在近亲之间进行的。我们将首先涉及已经考虑过的 4 种亲属之间的交配（其他的系统见 Wright，1933，1969）。然后我们将涉及由回交产生的近交以及一次杂交之后的世代中

的近交。

近亲近交

连续世代中的近交系数可从等式［5.4］到［5.7］中给出的共祖率中计算出来。但先得到将一个世代的近交系数和以前世代的近交系数关联起来的循环等式则更方便。我们感兴趣的世代用 t 表示，前一个世代用 $t-1$ 表示，前两个世代用 $t-2$ 表示；我们对这 4 个系统不得不向上追溯至 $t-3$ 世代那么远。循环等式的获得如下，并且连续世代中的近交系数在表 5-2 中给出。

自体受精 如果世代 t 的 X 是世代 $t-1$ 的 A 的后代，等式［5.4］给出：

$$F_X = f_{AA} = \frac{1}{2}(1 + F_A)$$

因此循环等式是：

$$F_t = \frac{1}{2}(1 + F_{t-1}) \qquad [5.8]$$

在第一个世代中，亲本是非近交的，$F_{t-1} = 0$，这使得 $F_{(t=1)} = \frac{1}{2}$。在第二个世代中，$F_{t-1} = \frac{1}{2}$，$F_{(t=2)}$变成$\frac{1}{2}\left(1 + \frac{1}{2}\right) = \frac{3}{4}$。以这种方式进行下去允许人们写下每一个连续世代中的近交系数。注意在这种情形中，近交率从一开始就是固定不变的，它与等式［3.7］完全一致：$\Delta F = \frac{1}{2N} = \frac{1}{2}$。对于其他系统，这不是正确的。自体受精借助一个正常的交配制度有可能产生最快速的近交。近交系数在 10 个世代之后可达到 99.9%。可是，在某一步骤中，通过一些形式的孤雌生殖以及通过诸如加倍单倍体细胞的染色体组的操作获得完全的纯合体是有可能的。

我们下一步将涉及全同胞交配，因为对于其他系统它是最经常被使用的。

全同胞 根据等式［5.6］的共祖率，参考图 5-5，有：

$$F_X = f_{PQ} = \frac{1}{4}(2f_{AB} + f_{AA} + f_{BB})$$

为了获得循环等式，我们必须将共祖率表示为以前世代的近交系数。首先，$f_{AB} = F_P = F_{t-1}$。因为在同一世代的个体有相同的近交系数，所以根据等式［5.4］有 $f_{AA} = f_{BB} = \frac{1}{2}(1 + F_A)$，并且 $F_A = F_{t-2}$。作出这些替代得循环等式：

$$F_t = \frac{1}{4}(1 + 2F_{t-1} + F_{t-2}) \qquad [5.9]$$

在第一个世代中，F_{t-1}和 F_{t-2}两者都为 0，因此 $F_{(t=1)} = 0.25$。在开始 4 个世代中的近交系数是 0.25，0.375，0.50 和 0.59。近交率在最初的几个世代中是不稳定的，这一点通过从等式［3.9］计算 ΔF 看出来。对最初的 4 个世代，ΔF 是 0.25，0.17，0.20 和 0.19。后来近交率达到一个固定值 0.191。根据等式［4.1］，于是有效群体大小是 $N_e = 2.6$。

后代—亲本 这里我们仅考察后代与它们较年轻的亲本的交配；与同一亲本反复回交将在以后考察。图 5-6 显示了和所需要的系谱一样多的系谱，字母与图 5-4 的相同，共祖率和等式［5.5］的相一致。每一个个体是一个世代的后代，并且是下一世代的一个亲本。

X 的近交由等式［5.5］给出，为 $F_X = f_{PA} = \frac{1}{2}(f_{AB} + f_{AA})$。通过代入 $f_{AB} = F_P = F_{t-1}$，和 $f_{AA} = \frac{1}{2}(1 + F_A) = \frac{1}{2} + \frac{1}{2}F_{t-2}$，可获得循环等式。于是，这个循环等式变成与等式［5.9］的全同胞循环等式完全相同。可是，这仅对常染色体基因是正确的；对性连锁基因，后代—亲本交配给出稍微较高的近交率，在开始几个世代之后 $\Delta F = 0.293$（Wright，1933）。

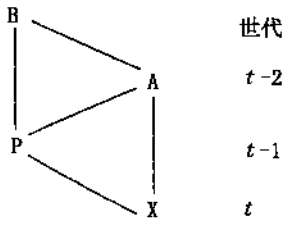


图 5-6

半同胞　图 5-1 给出了共祖率由等式［5.7］指明的个体。为了获得反复的半同胞交配的循环等式，我们必须知道个体 B 和 C 之间的亲缘关系。它们要么彼此是半同胞要么是全同胞。对于家畜，B 和 C 通常是雌性，两者都与同一雄性 A 交配。为了用总是半同胞的 B 和 C 的等价体继续半同胞交配，有必要使雌性中的一个与第二个雄性交配，这样在每一个世代中形成 4 个个体作为亲本。这在实践中是困难的，但如果它做到了，那么采用与上面一样的相同方式通过代入面获得的循环等式变成：

$$F_t = \frac{1}{8}(1 + 6F_{t-1} + F_{t-2}) \qquad [5.10]$$

用总是全同胞的 B 和 C 来继续半同胞交配更容易，于是每一世代中的亲本数为 3 个。于是近交变得较快一些，循环等式是：

$$F_t = \frac{1}{16}(3 + 8F_{t-1} + 4F_{t-2} + F_{t-3}) \qquad [5.11]$$

固　定

人们经常对因近交引起的固定概率比对近交系数更感兴趣。近交系数给出了一个个体是纯合体的概率，依表 3-1 它是 $1 - 2p_0q_0(1 - F)$。但是，人们还想知道一个系群中所有个体一特定位点上的同一个等位基因预计会多久达到纯合。固定概率取决于等位基因的数目以及它们在系群中起始交配中的排列。当 4 个等位基因出现在起始的交配中时，经过最初 20 个世代的全同胞交配，任何一个位点的固定概率给在表 5-2 的栏 B（2）中。当然，在一个同胞交配的系群中不可能有 4 个以上的等位基因，当有更少的等位基因时，固定概率就更大（见 Haldane，1955）。在起始几个世代之后，固定概率变得仅比近交系数小一点。

实际更重要的一个问题是所有位点上的固定概率或预计整个基因组被固定的比例。这是由术语“纯系”所暗示的“纯净”程度，“纯系”经常被用来指高度近交的系。我们不能从假定的位点数中得到总固定概率以及任何一个位点的固定概率，因为当一个位点变成固定时，与它连锁的邻近位点也变成固定了。因此，总固定概率取决于图距总长度，并且在较低的程度上，还取决于染色体数目。这是一个复杂的问题，这里不可能解释：细节见 Stam（1980）。总固定比任何一个位点的固定要花费长得多的时间，这一点再明显不过了。例如，一个象小鼠的生物体经过 20 个世代的同胞交配之后，实际上能预料到没有一个个体是完全纯合的，没有一个系被完全固定。预计 95% 的个体达到完全纯合将需要大约 50 个世代（Stam，1980）。在 50 个世代之后，总固定概率近乎相同。因此以相反的方式来表述这个问题，我们可得出结论：在小鼠的同胞交配系群中，预计经 50 个世代之后，约 5% 的基因组仍然是杂合的。由突变造成的新杂合性在第 15 章描述。

表 5-2 在各种近亲近交系统下的近交系数和全同胞交配下的固定概率

世代 (t)	A	B (1)	B (2)	C	D
0	0	0	0	0	0
1	0.500	0.250	0	0.125	0.250
2	0.750	0.375	0.063	0.219	0.375
3	0.875	0.500	0.172	0.305	0.438
4	0.938	0.594	0.293	0.381	0.469
5	0.969	0.672	0.409	0.449	0.484
6	0.984	0.734	0.512	0.509	0.492
7	0.992	0.785	0.601	0.563	0.496
8	0.996	0.826	0.675	0.611	0.498
9	0.998	0.859	0.736	0.654	0.499
10	0.999	0.886	0.785	0.691	
11		0.908	0.826	0.725	
12		0.926	0.859	0.755	
13		0.940	0.886	0.782	
14		0.951	0.908	0.806	
15		0.961	0.925	0.827	
16		0.968	0.940	0.846	
17		0.974	0.951	0.863	
18		0.979	0.960	0.878	
19		0.983	0.968	0.891	
20		0.986	0.975	0.903	

栏	交配系统	循环等式
A	自体受精或与高度近交系反复回交	$\frac{1}{2}(1+F_{t-1})$
B	全同胞的兄弟×姐妹，或后代×较年轻亲本：	$\frac{1}{4}(1+2F_{t-1}+F_{t-2})$
(1)	近交系数	
(2)	固定概率(来自Schäfer, 1937)	
C	半同胞(雌性半姐妹)	$\frac{1}{8}(1+6F_{t-1}+F_{t-2})$
D	与随机繁殖的个体反复回交	$\frac{1}{4}(1+2F_{t-1})$

反复回交

由于多种目的，经常进行与一个个体或与一个高度近交系的反复回交。引起的近交如下。系谱(图 5-7)显示一个个体 A，它可能是雄性，与他的女儿 C、他的孙女 D 等交配。根据补充规则[5.5]，

$$F_X = f_{AD} = \frac{1}{2}(f_{AA} + f_{AC})$$

$$= \frac{1}{2}\left[\frac{1}{2}(1 + F_A) + F_D\right]$$

因此循环等式是：

$$F_t = \frac{1}{4}(1 + F_A + 2F_{t-1}) \qquad [5.12]$$

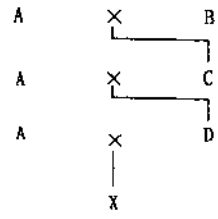


图 5-7

其中 F_A 是用来进行反复回交的个体近交系数。如果 A 是来自基础群体的一个个体，$F_A = 0$，则等式变成：

$$F_t = \frac{1}{4}(1 + 2F_{t-1}) \qquad [5.13]$$

经过最初 9 个世代的近交系数给在表 5-2 中。如果 A 是来自一个高度近交系的一个个体，并且 $F_A = 1$，则等式变成：

$$F_t = \frac{1}{2}(1 + F_{t-1}) \qquad [5.14]$$

这与自体受精的等式完全相同。在这种情形中，在连续世代中 A 不必是同一个个体：它可以是近交系的任何成员。

反复回交的主要用途是将一个特定的基因从一个品系转移到另一个品系的遗传背景中。于是出现了一个和期望基因一起被转移的外源染色体的长度问题。一个显性基因通过杂合体与它要被导入的品系连续杂交可被转移。可以看到（见 Crow 和 Kimura，1970，p.94）在这种情形下，经 t 次杂交后在该基因每一侧被导入的染色体平均长度大约是 $100/t$ cM，或总共是 $200/t$ cM〔1 厘摩（cM）相当 1% 重组频率的图距〕。一个隐性基因通常通过交替回交和互交来转移，从中可选取纯合体。在这种情形中，t 次循环之后，每一侧的外源染色体的平均长度大约是 $200/t$ cM，或者总共是 $400/t$ cM（Bartlett 和 Haldane，1935）。根据被转移的连锁染色体的长度以及该生物体的图距总长，我们能够获得总基因组仍然是杂合的期望比例。例如，假定一个显性基因经 5 次回交转移到一个近交小鼠品系中。借助该基因携带的连锁染色体的长度达到 200/5 = 40cM。取小鼠的图距总长为 1 600cM（Davisson 和 Roderick，1989），这种杂合的片段将占总基因组的 2.5%。另外，与被转移基因没有联系的基因组的某些部分仍然是杂合的。这可以近似地取为 $1 - F$，根据表 5-2 的 A 栏，在 5 次回交之后，它是 3%。因此预计总共大约 5.5% 的基因组仍然是杂合的。这个问题精确的处理由 Stam 和 Zeven（1981）给出。组织相容性基因的转移有特殊的问题，Johnson（1981）考虑到了这些问题。

杂交和后续世代

在遗传分析以及育种（特别是植物育种）中，标准的作法是在高度近交系之间进行杂交并产生 F_1、F_2 和后续的世代。如果这些被作为一个大的随机繁殖群体来保留，那么在后续世代中近交系数是多大？通过考察配子类型很容易回答这个问题，而且通过共祖率的规则来核对答案也没有困难。我们将考察从两个近交系的杂交（二元杂交）和 4 个近交系的杂交（四元杂交）中衍生出的群体。所有系群是完全近交的，因此任何一个系产生的所有配子都是同一的。因此，从二元杂交中衍生出的群体仅源于两个不同的配子；该群体等

价于从一个随机繁殖群体中由一个个体自体受精衍生出的群体，因此它的近交系数是 $F=0.50$。从四元杂交中衍生出的群体始于4个不同的配子，与全同胞交配所得群体相同，近交系数是 $F=0.25$。

衍生群体的这些近交系数没有意义，除非定义了它们涉及到的基础群体。在上面推理过程中暗示的基础群体是某个真实或假定的随机繁殖群体，近交系就从中获得。在杂交中使用的近交系被假定是平均基因频率没有任何变化，即没有选择，而产生的所有可能系群的一个随机样本。对用这种方式定义的基础群体，衍生群体的近交系数的意义如下述。如果我们进行大量的二元或四元杂交，各有不同的近交系组，那么从杂交中衍生的群体将构成一组系群或亚群体。于是近交系数将表明在这些系群间基因频率离散的期望数值。从二元杂交中衍生的群体等价于自体受精的一个世代的后裔。因此，基因频率只能有3个值：0、$\frac{1}{2}$和1。从四元杂交中衍生的群体等价于全同胞交配的一个世代的后裔，基因频率只能有5个值：0、$\frac{1}{4}$、$\frac{1}{2}$、$\frac{3}{4}$和1。

混合的近交和杂交

很多植物是“近交者”，通过自体受精正常繁殖。可是，这些植物中有很多植物，也经常发生一些异花授粉。杂交的比例变化很广泛，例如在青豆和高粱品种中，变化范围为5%～50%（Allard，Jain 和 Workman，1968）。杂交将产生多大的杂合性？假定整个群体很大，一个个体是自交还是杂交都是随机的，与它的双亲进行的一切无关。在任何一个世代，有两种后裔，由自体受精产生的后裔以及由杂交产生的后裔。让 C 表示由杂交产生的个体比例；它们的近交系数为0。由自交产生的比例是 $(1-C)$，按等式［5.8］，它们的近交系数是 $F_t=\frac{1}{2}(1+F_{t-1})$。因此，平均近交系数是：

$$F_t=\frac{1}{2}(1+F_{t-1})(1-C)$$

如果杂交速率保持固定，平均近交系数就会达到一个平衡水平，它在这个水平上保持不变。于是 $F_t=F_{t-1}$，对上面的等式重新整理给出平衡时的平均近交系数为：

$$F=\frac{1-C}{1+C} \qquad [5.15]$$

其中，C 是由异花授粉产生的个体比例。根据没有进行有利于或不利于杂合体的选择这一假设，由等式［3.15］可知，相对于完全随机交配群体的杂合体期望频率是 $1-F$。等式［5.15］和［3.15］的应用表明，5%杂交产生的杂合性等于随机交配群体杂合性的9.5%，50%杂交产生的等于66.7%。对大麦的研究已经表明，4个酯酶位点上杂合体的频率大于由已知杂交数量所预测的值，它是0.57%，过量的杂合体归因于有利于杂合体的选择（Allard，Kahler 和 Weir，1972）。

杂交对一个主要近交群体的结构影响比对杂合性产生的影响更重要。没有杂交，一个近交群体完全由纯合系群组成，自然选择通过对不能很好适应的系群的淘汰而起作用。于是每一个当地栖息地都被最适应它的系群所居住（Allard，Jain 和 Workman，1968），但不会发生进一步的适应，对不同栖息地的新适应也不会发生，除非有突变。可是，借助一些杂交，由于现存系群的基因重组，新的系群可以不断地产生，这允许连续的或新的适应性

发生。杂交还使得淘汰因突变引起并因近交固定的有害基因成为可能。

相反的问题，即在主要远交群体中的小数量的近交，也是有趣的。将 C（杂交比例）的高值代入等式［5.15］表明小数量的自交造成的平均近交系数增加极小。其理由是群体没有分化成永久的系群；自交的后裔很像是杂交繁殖的后裔。如果近交是通过全同胞交配而不是自交进行的，那么用类似于等式［5.15］的方式表示的近交系数是 $F=(1-C)/(1+3C)$。以同胞交配产生的个体比例 $S=1-C$ 的形式表示，公式变成：

$$F=\frac{S}{4-3S} \qquad [5.16]$$

（细节见 Li，1976）。例如，一个大群体中 5% 的全同胞交配可将平均近交系数从 0 增加到 1.2%，10% 的全同胞交配可将它增加到 2.3%。这个情形对实践的启示是，通过随机交配或通过最小的近交来保持一个种群，不必担心偶然的同胞交配的后果，这一点在第 4 章已注意到了。

基础的改变：结构化群体

这里要考察的问题不限于系谱群体或近亲近交。对于确定为基础群体的一组个体，已经计算了它的近交系数，于是人们可能想知道不同基础的近交系数，无论其祖先或是很遥远或是不很遥远。例如，由全同胞交配产生的个体有 25% 的近交涉及其双亲。其双亲本身可能涉及是近交的更遥远的基础。涉及这种更遥远的基础群体的个体近交系数是多大？这个问题暗示了具有等级细分成系群和亚系群的一种“结构化”群体，如图 5-8 所示。在图 5-8 中，A 代表我们所关注的进一步向上追溯的基础群体，B 是稍后一个阶段的基础群体，X 代表要计算近交系数的个体。与 B 同时期的无字母的圆圈代表以图 3-1 的方式将 A 细分的各个系群。在真实群体中，这些系群中只有一个 B 可能是实际存在的。于是系群 B 进一步分化成亚系群，X 是这些亚系群之一的一个个体。从杂合体相对频率来求解答案。让 H_X、H_B、H_A 分别表示 X、B 和 A 的同时代个体之间的杂合体频率。于是 $H_X/H_A=(H_X/H_B)(H_B/H_A)$，根据等式［3.15］，它是：

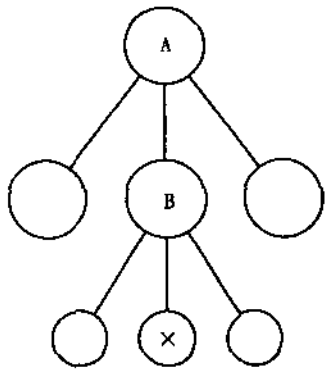


图 5-8

$$P_{X\cdot A}=P_{X\cdot B}P_{B\cdot A} \qquad [5.17]$$

其中 $P_{X\cdot A}=1-F_{X\cdot A}$；$F_{X\cdot A}$是 X 以 A 作为基础的近交系数，其他的下标也类似。等式［5.17］的关系能够扩展到任何数目的细分类别或任何的近交阶段。

等式［5.17］与用来描述结构化群体的被称为 *Wright* 氏 F－统计量的表达方式相同。采用 F-统计的术语，F_{IS}是个体相对于它自己的亚群体的近交系数，与上面的 $1-P_{X\cdot B}$相等；F_{ST}是亚群体相对于整个群体的平均近交，与上面的 $1-P_{B\cdot A}$相等；F_{IT}是相对于整个群体的个体近交系数。等价于等式［5.17］的关系是：

$$(1-F_{IT})=(1-F_{IS})(1-F_{ST})$$

例5.2

小鼠的一个品系繁殖了42个世代，其有效群体大小约为40，然后又通过11个世代的全同胞交配繁殖（Falconer，1971）。最后近交系数是多大？根据表5-2，由全同胞交配即图5-8中从B到X产生的近交是0.908。因此 $P_{X\cdot B}=1-0.908=0.092$。当同胞交配开始，即从A到B时，系群的近交如下：由于 $N_e=40$，等式［4.1］给出 $\Delta F=0.0125$，经过42个世代之后，等式［3.12］给出 $F_{B\cdot A}=0.410$。因此 $P_{B\cdot A}=1-0.410=0.590$。根据等式［5.17］，$P_{X\cdot A}=0.590\times 0.092=0.054$。因此，以系群的起源作为基础，最后的近交系数是 $F_{X\cdot A}=1-P_{X\cdot A}=0.946$。

突　变

在长时期的近交之后，突变可能成为确定杂合体频率的一个重要的因素。如果 u 是系群中已达到接近固定的一个基因的突变率，那么在自体受精下对于常染色体位点，这个位点上由突变引起的杂合体频率是 $4u$，在全同胞交配下是 $12u$（Haldane，1936）。如果我们只关注一个位点，这些是很小的频率，但如果所有位点的效应加在一起，那么突变作为长期近交品系（如广泛使用的小鼠品系）中杂合性的一种来源是不能完全忽略的。由突变引起的异质性来源的实际后果是一个系群的特性通过突变等位基因的固定而慢慢改变，并且亚系群变成分化的了。这个问题在第15章中进一步考察。

有利于杂合体的选择

当实施近亲近交时，目标一般是要产生系群内的固定或纯合性。因此了解选择将如何影响趋向固定的进程是一个具有某种重要性的问题。对一个有害隐性等位基因不利的选择可能阻止有害等位基因被固定，但它不能延缓较有利等位基因的固定。可是，有利于杂合体的选择是另一个问题。几乎普遍观察到近交的一种后果是适合度下降，其理由将在第14章给出。因此选择抵抗近交，因为较纯合的个体适合度低，这只能意味着选择有利于杂合体——不必是单个位点上的杂合体，而是染色体片段上的杂合体。两个有害基因（隐性的或部分隐性的）相斥连锁是必需的，借以给两个基因所在染色体片段上杂合体一个选择优势。因此，为了评价计算的近交系数作为固定概率一种度量的可靠性，探讨近交和有利于杂合体的选择这两种对抗的趋势如何彼此平衡是重要的。

近交和有利于杂合体选择的联合作用的结果取决于是否有较适合的系群来更新较不适合的系群；换句话说，取决于选择是在系群之间还是仅在系群内操作。在任何一个系群内，不利于纯合体的选择仅延缓趋向固定的进程，但不能阻止它，除非两个纯合体完全不能生存或完全不育。延缓大致与不利于纯合体的选择系数成比例，它在最亲近的近交下最小（Reeve，1955*a*）。例如，如果选择同等地不利于两个纯合体，选择系数 $s=0.2$ 时，则在自交不育品系中近交的“实际”速率 F 将从0.50减少到0.44；$s=0.75$ 时则减少到0.20。对全同胞交配，这些选择系数分别将 F 从0.19减少到0.15和0.03，所有这些值只是近似的。果蝇（*Drosophila*）的一个实验提供了延缓近交的一个例子（Rumball等，1994）。两个同胞交配系群和由双重表兄妹近交产生的两个系群分别持续了11个和18个世代，从6个酶位点的基因型频率中估计了实际的近交。所有4个系群中近交的实际速率大约是理论速率的80%，延缓可明确地归

因于有利于杂合染色体片段的自然选择。

可是，如果存在很多系群并且交配系统有利于较不适合的系群被来自较适合系群的亚系群所更新，则结果有所不同。例如，如果混合收集自花传粉植物群体的种子并作为一个随机样本用于播种；或者在一个同胞交配群体中，如果交配的一对同胞随机取自所有存活的后代，则将发生更新。处于任何阶段的系群在留下的杂合体数目上会偶然不同，具有较多杂合体的系群将留下更多的后代，因而对下一世代贡献最大。对于这种繁殖系统，不利于纯合体的选择，如果足够强烈，可导致近交和选择相反力量之间的一种平衡状态，离散过程被阻止。阻止趋向于纯合性所需要的选择系数是相当高的。例如，如果两个纯合体被同等地进行不利选择，则阻止完全纯合性需要的最小选择系数，在自体受精群体中是 $s = 0.50$，在全同胞交配群体中为 $s = 0.24$。当杂合体的频率是它起始值的一半时，即 $F = 0.5$ 时，终止这个过程的选择系数对自体受精为 $s = 0.67$，对全同胞交配为 $s = 0.47$（Hayman 和 Mather，1953）。

上面概述的结论指的是一单个位点。如果在不同的染色体上具有几个位点，所有位点都经受强度足以阻止或严重延缓近交进程的不利于纯合体的选择，那么来自所有位点的适合度的总损失将是很严重的。具有高繁殖率的生物体（例如植物和果蝇）的近交系可能很好地经得起适合度的总损失，足以维持几个位点或染色体片段永久不被固定。但在一个如小鼠的生物体中，阻止两个或三个以上位点被固定包含的适合度损失将被削弱。在实验室条件下，小鼠的高度近交系，经 100 或更多世代的同胞交配之后，具有的适合度还不及非近交系的一半。可以想象它们可能有一个永久未被固定的位点，但很难相信它们能有更多的这样位点。一个位点上两个纯合体完全致死或不育意味着失去了 50% 的后代；在两个不连锁的位点上，则失去了 75%。具有 50% 死亡率或不育性的一个小鼠品系可以繁殖下去，但如达 75% 时就很难繁殖了。

习　题

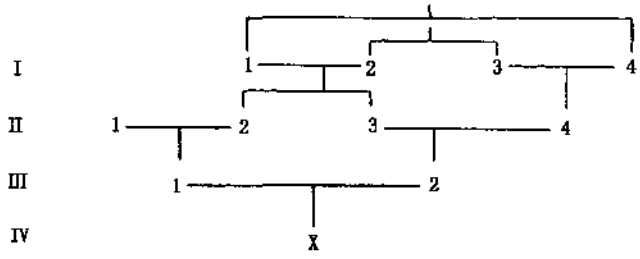
5.1　下面亲属之间结婚产生的后代中近交系数是多大？（1）单重亲表兄妹，(2) 双重亲表兄妹，(3) 叔父—侄女。

［解答序号：4］

5.2　一对同卵双生子和不相关个体结婚产生的孩子的共祖率是多大？

［解答序号：14］

5.3　下面是没有胼胝体的人类系谱。在世代Ⅰ中，个体 1 和 4 是全同胞，2 和 3 也是全同胞。在世代Ⅳ中，X 代表 8 个个体中有 2 个患病个体的一个家庭。计算这个家庭的近交系数以及它的亲本Ⅲ2 的近交系数。



数据来自 Shapira，Y.& Cohen，T.（1973）*J. Med. Genet.*，**10**，266～269。

［解答序号：24］

5.4 如果一个以自花受精为主的植物正常地具有 1% 频率的异花授粉，假定没有选择，在基因频率为 0.2 和 0.8 时，两等位基因位点上的杂合体频率将是多大？

［解答序号：34］

5.5 假定以自花受精为主的植物群体有两个等位基因 a 和 b 是多态性的，且 3 种基因型的频率是：

aa	ab	bb
0.54	0.12	0.34

假定没有选择，这表明异花授粉的频率有多大？

［解答序号：44］

5.6 有效大小为 16 的一个群体保持 10 个世代，其近交系数是多大，具有加倍的有效大小时再保持 10 个世代，那么其近交系数又是多大？它与多少个世代的全同胞交配相当？

［解答序号：54］

5.7 一种植物的两个高度近交系杂交产生 F_1 世代。F_1 个体自交产生 F_2 世代。然后 F_2 个体与 F_1 以及近交系群之一回交。这两种回交的后代中近交系数是多大？

［解答序号：64］

5.8 同血缘的婚姻增加了孩子患隐性疾病的风险。计算出在下列两种情形下，这种风险会被堂兄妹婚姻（单重亲堂兄妹）增加多少：（1）具有群体发病率为1/2 500的囊性纤维变性，（2）发病率为1/11 000的苯丙酮尿症。

［解答序号：74］

5.9 假定群体中个体的一个比例 y 是由血缘交配产生的，这使得这些个体都具有近交系数 F，而剩余的 $1-y$ 是由随机交配产生的，例如，具有一些堂兄妹婚姻的人类群体。如果一个隐性基因的纯合体在整个群体中的发生率为 I，证明这个隐性等位基因的基因频率 q 可通过下式从整个发生率 I 中估计出来：

$$(1-yF)q^2 + yFq = I$$

［解答序号：84］

第 6 章　连续变异

生物学家以及外行人都会明显看出，前几章讨论的那类变异只包括一小部分自然发生的变异。一个人只要环顾一下周围的男男女女，便能完全认知他们在诸多方面是完全不同的，这些差异几乎都是程度上的不同，很少存在可归因于单个基因分离的明显区别。例如，如果我们按照个体的高度来对个体分类，我们不能将它们分到标有“高”和“矮”的两组，因为有不同程度的高度，分类的分界线完全是武断的。这类没有自然间断性的变异，称之为*连续变异*，展现这类变异的性状，称之为*数量性状*或*计量性状*，因为对它们的研究取决于度量而不是计数。作为数量性状遗传基础的遗传学原理，基本上就是前面几章概述的*群体遗传学*的那些原理。但是由于有关基因的分离不能逐个研究，因此需要新的研究方法以及必须引入新的概念。关于数量性状的遗传学分支称为*数量遗传学*或*统计遗传学*。这个遗传学分支的重要性已不必强调了；对植物和动物育种者而言，具有经济价值的大多数性状都是数量性状，与微观进化有关的大多数变化也是数量性状的变化。因此正是在这个分支中，遗传学才对实际问题有最重要的应用，也才对进化理论有最直接的影响。

由遗传分离引起的内在不连续变异转化为数量性状的连续变异是如何实现的呢？有两个原因：一是影响性状的很多基因同时分离，另一是由非遗传原因引起的真实连续变异的叠加。例如，考察一种简化情形。假设在 6 个未连锁位点上发生分离，每个位点有两个频率为 0.5 的等位基因。假设每一个位点有一个完全显性的等位基因，并且每一个显性等位基因令指定性状的度量值增加一个单位。于是如果这些基因的分离是变异的唯一原因，按照个体是否有显性等位基因出现在 0、1、2、…或 6 个位点上，则在性状的度量值中将有 7 个不连续的类别。这些类别的频率可由二项展开式$\left(\frac{1}{4}+\frac{3}{4}\right)^6$求出，如图 6-1（*a*）所示。如果我们的度量足够精确，我们应当清楚地识别这些类别，而且我们应当能够含义明确地将任何一个个体归入它所属的类别。如果有更多基因分离，而且每一基因具有较小的效应，那么将有更多的带有它们之间更小差异的类别，如图 6-1（*b*）所示。于是区别这些类别将更加困难，如果这些类别之间的差异变得和度量误差大致一样小，则我们不再能识别这些不连续性。另外，数量性状含有非遗传原因的变异，这种变异是真正的连续变异。无论我们的度量多么精确，这种变异的效应似乎使遗传不连续性的界线模糊不清，以致于我们看到的不连续变异成为连续的。

因此，有关孟德尔性状的基因与有关数量性状的基因之间的区别在于同变异的其他来源相比它们的效应大小。一个基因其效应足够大，甚至在其他位点上的分离和非遗传变异都存在的情况下也能引起可识别的不连续性，这种基因能用孟德尔方法研究，而效应不足以引起不连续性的基因，便不能单个研究。术语*主效基因*和*微效基因*反映出这种区别。然而存在着很多中间级基因，它们不能被准确分为主效基因或微效基因。而且，作为基因多效性的结果，同一基因对一个性状可归类为主效，而对另一性状却是微效的。因此这种区

分，虽然很方便，但不是基本的，没有充分的证据表明有两类不同性质的基因。由很多基因同时分离造成的变异可称为多基因的变异，与之有关的微效基因有时也称为多基因。

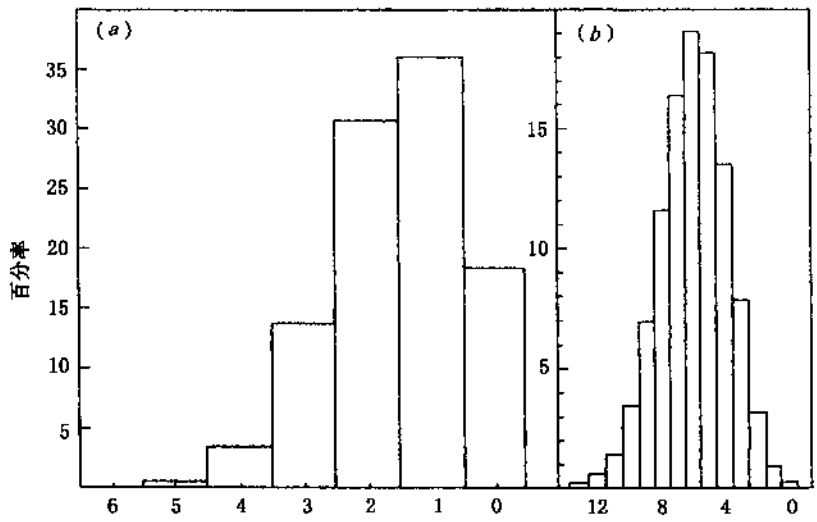


图 6-1　几个或很多位点中每一个位点上两个等位基因同时分离的期望分布

(a) 6个位点，(b) 24个位点。每一个位点上一个等位基因对另一个为完全显性，基因频率都是0.5。当隐性等位基因纯合时，假定每一个位点减少度量值，(a) 中为1个单位，(b) 中为$\frac{1}{4}$个单位；表示度量值的水平尺度指出了隐性等位基因纯合的位点数，垂直轴指出每一类别中期望个体的概率或百分率。概率可由二项展开式$\left(\frac{1}{4}+\frac{3}{4}\right)^n$获得，其中 n 是位点数。

数量性状

在任何高等生物体中可研究的数量性状几乎是无限的。连续变化并能度量的任何属性，大体上都可以作为数量性状来研究，例如，解剖上的大小和比例、各种生理机能、精神或心理素质。基本条件是这些属性应当可以度量。可是，度量方法为我们怎样研究设置了实际限制。通常有大量的个体必须度量，然而对需要用精心设计的技术去度量的任何性状研究就变得行不通了。因而在数量遗传学研究中使用的性状主要是解剖上的大小、或者按照终产物度量的生理机能，例如泌乳、生殖力或生长速度。

图 6-2 说明了数量性状的一些例子。根据度量的频率分布以图表形式表示变异。度量结果分成间隔相等的类别，以垂直尺度绘出归入每一类别的个体比例。仅为了绘制的方便，由此得出的柱形图不连续。如果类别范围变得较小并且度量的个体数无限地增加，则柱形图将变为一条平滑曲线。一些像刚毛数或窝产仔数的数量性状变异，严格地说不是连续的，因为它们是通过计数而被度量的，它们的值只能是整数。尽管如此，根据类似于将度量值分成类别的方法，我们可以认为这些情形下的度量结果起源于潜在的性状，而这种潜在性状虽然只能用整数表示，但其真正变异是连续的。例如，窝产仔数可以认为是潜在的、连续变异的性状——生殖力的一种度量。出于实用目的，只要类别数目不太少，这类性状可以作为连续变异来处理。当类别数太少时，例如以死亡或存活表示疾病易感性时，

必须使用不同的方法，这一点将在第 18 章解释。

大多数数量性状的频率分布或多或少接近于正态曲线。这一点可以从图 6-2 中看出，其中用柱形图绘制的平滑曲线是正态曲线，其平均数和方差从数据中计算得来。因此在数量性状的研究中，利用正态分布的特性和应用适宜的统计方法是可能的。可是，如果要获得接近正态的分布，有时必须修改度量的尺度。例如，如果简单地以小眼面数目来度量和绘图，则图 6-2（*d*）的分布将是偏斜的。但如果以对数单位度量和绘图，则它会变为对称并接近正态分布。选择度量尺度所依赖的标准在这里不能全面地叙述，将在第 17 章解释。同时假设讨论中的任一数量性状都以适宜尺度度量并且具有一个近似正态的分布。

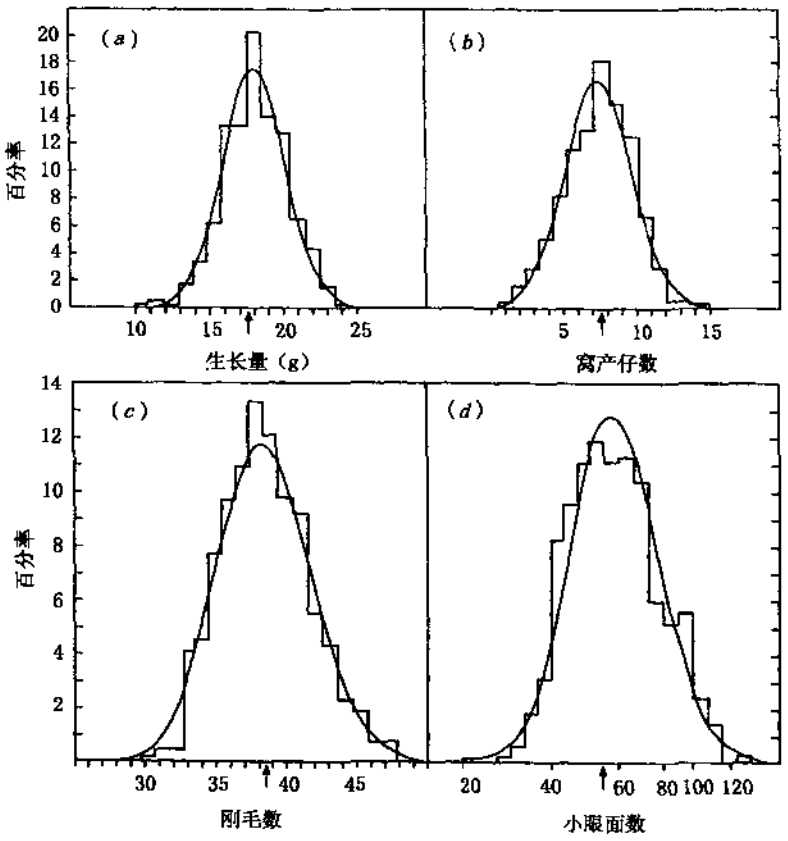


图 6-2 4 个数量性状的频率分布，具有附加的正态曲线

平均数用箭头标明。性状如下，每一柱形图基于的观察数目在括号中给出：

（*a*）小鼠（♂♂）：3～6 周龄的生长。(380)

（*b*）小鼠：窝产仔数（第 1 窝中活幼鼠数）。(689)

（*c*）黑腹果蝇（♀♀）：第 4 和第 5 腹部片段的腹面上刚毛总数。(900)

（*d*）黑腹果蝇（♀♀）：突变体“棒”眼中小眼面数。(488)

（*a*）、（*b*）和（*c*）来自原始数据；（*d*）来自 Zeleny（1922）的数据。

认为数量性状实际上由很多位点上的基因所控制的根据是什么？有两类证据。第一类是间接的：观察值与依据多基因遗传理论得到的期望值普遍吻合得很好。第二类是直接的

实验证据：影响数量性状的位点通过第 21 章描述的方法可以定位在染色体上。以这种方法识别的位点称为数量性状位点或 QTL。黑腹果蝇（*Drosophila melanogaster*）在胸部和腹部有可变数目的感觉刚毛，分别称为腹侧刚毛和腹部刚毛。至少有 17 个影响腹侧刚毛数的 QTL 和 5 个影响腹部刚毛数的 QTL 已被定位在第 3 号染色体上（Shrimpton 和 Robertson，1988*b*；Long 等，1995）。类似地，已发现至少有 6 个影响果实产量的 QTL 可用来将栽培番茄与亲缘野生种鉴别开来（Paterson 等，1988）。

数量性状的特性

有两个基本遗传现象与数量性状有关，所有生物学家或多或少熟悉它们，每一种现象形成一种育种方法的基础。第一个是亲属间的相似性。每一个人都熟悉这个事实，亲属彼此倾向于很相像，一般地说亲缘关系越近，相似性就越大。虽然只在我们人类中这种相似性不用度量就很容易辨明，但这种现象在其他物种中同样存在。相似度随性状而变化，一些性状较高，一些较低。亲子之间的相似性提供了选择育种的基础。用较理想的个体作亲本可引起下一世代平均水平的提高，如同一些性状比另一些性状显示出较强的相似性一样，一些性状比另一些性状对选择更易于起反应。亲属之间的相似度是能够容易观察到的群体特性之一，它也是数量遗传学的目的之一，以便说明不同种类亲属间的什么样的相似性程度可用于预测选择育种的结果并指明实施选择的最佳方法。遗传学原理应用于家畜的选择育种已经获得了相当巨大的改进，产生了巨大经济效益；获得的改良速度每年大约在 1% ~ 3% 之间（Smith，1984，1988）。选择育种的基础原理形成了以下 7 章的中心题目，第 9 章和第 10 章阐述亲属间的相似性，第 11 ~ 13 章阐述选择效果。

很多数量性状与其他数量性状相关。例如在小鼠中，例示在图 6-2 中的两个性状，身体大小和窝产仔数相关；较大的小鼠倾向于有较多的窝仔数，正如在图 6-3 中所看到的那样。类似这样的相关部分是由基因多效性引起的；影响一个性状的一些基因也影响另一个性状。因此，一个相关性状可以用来辅助旨在改变与它相关的那个性状的选择。相关性状在第 19 章中涉及。

第二个基本遗传现象是近交衰退以及与它相反的杂种优势。外行人对这种现象不及对第一种熟悉，因为禁止乱伦的法律阻止了它在我们人类中较明显的显示；但动植物育种者对它非常熟悉。近交倾向于降低与动物和自然远交植物适合度紧密相关的所有性状的平均水平，结果导致了整体生活力和生殖力的下降。由于家畜和植物中具有经济价值的大多数性状都是生活力或生殖力的组成部分，因此一般来说近交是有害的。近交系下降的生活力和生殖力可通过杂交来恢复，在特定条件下这种杂种优势可用来作为改良的一种方法。商用种植的玉米产量的巨大提高就是用这种方法获得的，美国自 1935 年以来玉米产量大约增加了一倍（Hallauer 和 Miranda，1981，p.408；一些其他作物的改良见 Simmonds，1979，p.47）。近交和杂交的效果将在第 14 ~ 16 章中描述。

我们能观察到的与数量性状有关的群体属性是平均数、方差和协方差。群体自然细分为家系有利于我们分析方差的各种组分，这些组分形成亲属间相似度度量的基础。另外我们也能观察实验上应用诸如选择、近交或杂交育种等育种方法的结果。数量遗传学实践上的目的就是探知如何利用在目前群体中获得的观察结果来预测任一特定育种方法的结果。更一般的目的是探知观察到的群体属性是如何受有关基因的属性以及可能影响数量性状的

各种非遗传因素的影响。必须考虑的基因主要属性是显性度、不同位点的基因与其效应相联合的方式、基因多效性、连锁和自然选择下的适合度。要同时考虑所有这些属性，还有各种非遗传因素，将使问题变得复杂化而不能处理。因此，我们必须从较简单情形开始，通过一次处理一件事情来简化问题。

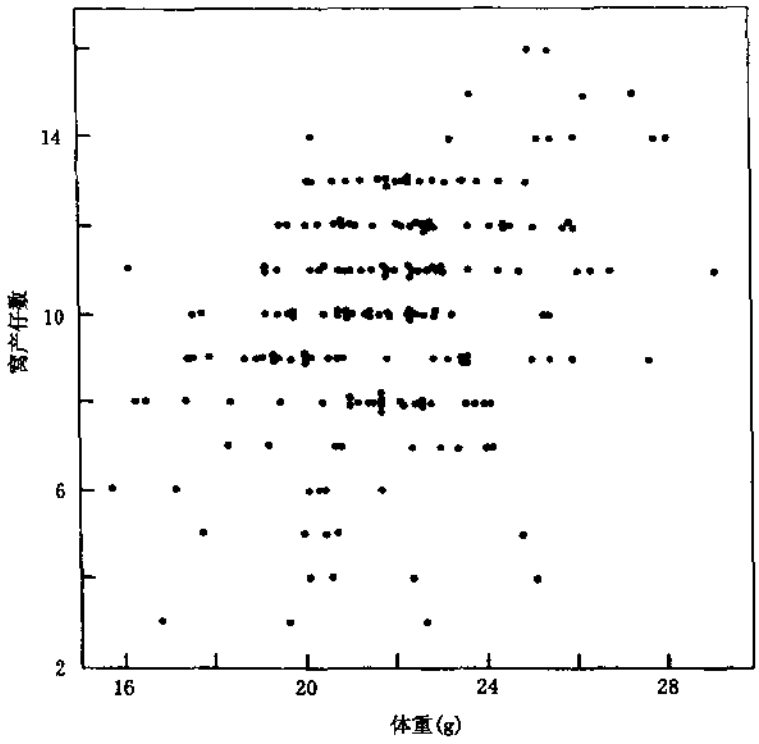


图 6-3　小鼠体重和窝产仔数之间的相关

每一点代表按照其 6 周龄体重和出生几周之后窝中活仔鼠数绘图的一单个雌性小鼠。这 200 个个体之间的相关是 0.3（Ian Hastings 博士热心提供数据）。

以下连续几章中紧跟着的计划是这样的：首先我们将看看什么决定群体平均数，然后引入两个新概念——平均效应和育种值——它们对理解方差是必需的。然后我们将讨论方差及其组分以及亲属间的协方差，这将引导我们了解亲属间的相似程度。在所有这些讨论中，我们一开始就充分考虑显性，对涉及到的其余复杂因素则简单地讨论一下。我们将作出的最重要的简化涉及基因对适合度的影响：我们将假定孟德尔分离不受基因型有差异的适合度的干扰。平均数、方差和协方差的描述将涉及一个随机交配群体，该群体具有哈迪—温伯格平衡基因型频率，没有选择，也没有近交。也就是说，我们将描述在任何特定育种方法应用之前的那个群体。于是在第 11 ~ 13 章中，我们将描述选择的效果，在第 14 ~ 16 章中则描述近交的效果。这将涵盖数量遗传学的基本原理，在最后几章中我们将讨论一些专题，包括第 19 章的相关性状、第 20 章的自然群体和进化、第 21 章的控制数量性状的基因本性。

习 题

6.1 下表列出的数字是来自两个栽培烟草品种杂交的 25 株 F_1 代和 25 株 F_2 代植株中每株的叶片数目，这两个品种的平均数分别为 15.0 片叶和 17.9 片叶。将 F_1 和 F_2 世代的频率分布制成表（如果愿意也可绘图）。从每一个分布中计算平均数、方差以及平均数的标准误。两个分布之间的主要差别是什么？

F_1					F_2				
18	15	16	18	15	16	20	19	17	14
16	14	16	18	17	16	14	14	15	17
16	13	16	14	16	20	13	12	15	16
15	16	15	15	16	21	18	15	14	18
15	16	16	15	16	14	17	13	15	13

数据来自 Johnson J.（1919）*Genetics*，**4**，307～340。

[解答序号：5]

6.2 下表给出了 50 只雄性小鼠 60 日龄的体重（g），它们由 12 对亲本繁殖而来，这些亲本处于选择小体重而育成的一个品系的一个世代中。将频率分布制成表（如果愿意也可绘图）并计算平均数和方差。分布的特点是什么？

12.8	15.7	13.7	5.7	6.1	13.2	12.5	7.0	6.6	14.8
12.8	14.1	14.0	17.8	14.7	11.8	13.5	14.1	16.2	6.6
5.7	11.7	13.4	6.6	6.8	15.6	13.0	11.5	12.1	15.4
11.8	11.4	10.7	15.6	14.9	4.7	13.1	13.4	14.5	13.4
11.8	17.2	15.0	15.1	14.6	14.8	15.1	14.9	18.4	16.5

J.W.B.King 博士热心提供数据。

[解答序号：15]

6.3 一个数量性状受 4 个独立分离的位点影响，每个位点有 2 个等位基因并且是完全显性，隐性等位基因纯合时每个基因减少 1 个单位的度量值，所有隐性等位基因的频率均为 0.3，据此计算基因型类别的频率分布，并根据频率分布来构建一个柱形图（像图 6-1 一样）。

[解答序号：25]

6.4 除两个位点上隐性等位基因频率为 0.3 和其余两个位点频率为 0.7 外，所有其他条件与习题 6.3 的相同，计算频率分布。

[解答序号：35]

6.5 有3个位点，每一位点有2个等位基因且无显性。在所有位点上，纯合体有2个度量单位的差异，杂合体与两种纯合体均有1个单位的差异。所有位点上降低度量值的等位基因的基因频率为0.4，据此计算频率分布。

[解答序号：45]

6.6 在（1）习题6.3和（2）习题6.5的条件下，什么基因频率将产生度量值类别的完全对称分布？

[解答序号：55]

第 7 章　数值与均数

在前几章中我们已经看到群体的遗传特性可以采用基因频率和基因型频率的形式来表示。为了推导基因和基因型频率为一方与数量性状显示出的数量差异为另一方之间的联系，我们必须引入一个新的概念，即*数值*的概念，它可以用数量单位表示，通过它性状得以度量。当在一个个体上度量性状时，观察到的数值是该个体的*表型值*。所有的观察值，不管是平均值、方差还是协方差，显然都必须以表型值的度量为基础。为了分析群体的遗传特性，我们必须将表型值分成归因于不同原因的组分。对这些组分含义的解释是这一章中我们主要的关注点，虽然我们还能够获知群体平均值是如何受基因频率阵列影响的。

表型值首先分剖成归因于基因型和环境影响的组分。*基因型*是个体具有的基因的特定集合，*环境*是影响表型值的所有非遗传因素。所有非遗传因素都包含在术语“环境”内，意味着基因型和环境，在定义上是表型值的唯一决定要素。与基因型和环境相联系的数值的两个组分是*基因型值*和*环境离差*。我们可以认为基因型给予个体一确定值，而环境在一个方向或另一个方向上引起偏离这个值的一个离差。或者，用符号表示为：

$$P = G + E \qquad [7.1]$$

其中 P 是表型值，G 是基因型值，E 是环境离差。整个群体中平均环境离差被认为是 0，因此表型均值等于基因型均值。于是术语*群体平均值*相等地适用于表型值或基因型值。当处理连续世代时，悟氮讥起见，我们假定环境在世代间保持固定不变，因此在没有遗传变化的情况下，群体平均值固定不变。如果我们能在大量的个体中复制一特定基因型，并在对群体是正常的环境条件下度量它们，那么它们的平均环境离差将是 0，因此它们的表型均值将等于该特定基因型的基因型值。这就是一个个体的基因型值的含义。原理上它是可度量的，但实践中它是不可度量的，当我们关注可为表型区分的一个单个位点上的基因型，或者高度近交系表现的基因型时除外。

为推论的目的，我们必须给讨论中的基因型指定任意的数值。可采用下面的方法来做到这一点。考察具有两个等位基因 A_1 和 A_2 的一单个位点，我们把一个纯合体的基因型值定为 $+a$，另一个纯合体的基因型值定为 $-a$，杂合体的基因型值定为 d（我们将采用 A_1 是使数值增加的等位基因这一习惯做法）。因此我们有像图 7-1 那样的基因型值的尺度。在这个尺度上原点或 0 值的那个点处于两个纯合体值之间的中点。杂合体的值 d 取决于显性度。如果无显性，$d = 0$；如果 A_1 对 A_2 是显性，则 d 是正的，如果 A_2 对 A_1 是显性，d 是负的。如果显性是完全的，d 等于 $+a$ 或 $-a$，如果存在超显性，则 d 大于 $+a$ 或小于 $-a$。显性度可表达为 d/a。

例 7.1

为本章和以后说明的目的，我们将提到 King（1950，1955）以及 Warwick 和 Lewis（1954）描述的小鼠的一种矮化基因，称为“侏儒”（符号 pg）。这个基因缩小身体大小，

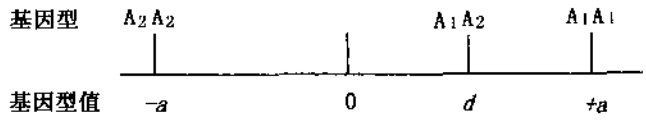


图 7-1　任意指定的基因型值

它对大小的影响效应几乎是（但不完全是）隐性的。在上面引用的研究进行到某个时刻，它存在于一个小的小鼠品系(MacArthur's)中。表 7-1 是小鼠 3 种基因型的 6 周龄体重，大约如下（两性平均值）：

表 7-1　小鼠 3 种基因型 6 周龄体重

	基因型		
	+ +	+ pg	pg pg
体重（g）	14	12	6

（这里给出的杂合体体重在某种程度上是推测的，但是误差不可能大于 1g。）这些是在正常环境条件下获得的平均体重，因此它们是基因型值。两个纯合体之间基因型值的中点是 10g，在像图 7-1 那样指定值的尺度上，这是原点，或零点。因此，在这种尺度上 a 的值是 4g，d 的值是 2g。

群体平均值

现在我们能够看看基因频率如何影响整个群体中性状的平均值。假定 A_1 和 A_2 的基因频率分别为 p 和 q。于是，表 7-2 的头两栏显示了 3 种基因型以及在一个随机交配群体中由公式［1.2］得来的它们频率。第三栏显示了它们的基因型值，像上面指定的那样。整个群体的平均值通过将每一种基因型值乘以它的频率并对 3 种基因型求和而获得。这为什么产生了平均值，其理由可通过将频率转换成个体数来理解。将数值乘以每一种基因型的个体数，并对基因型求和就给出了所有个体值的总和。于是平均值将是这种值的总和除以个体总数。在用频率进行操作的过程中方法是相同的，但因为频率的总和为 1，所以数值×频率的和就是平均值。换句话说，在获得频率过程中除以总数已经进行过了。用数值乘以频率来获得平均值是在本章和随后各章中经常用到的一种方法。回到群体平均值上，数值乘以每一种基因型的频率显示在表 7-2 的最后一栏。通过注意 $p^2 - q^2 = (p+q)(p-q) = p - q$，这一栏的总和可被简化。因此，群体平均值，即是这一栏的总和，为：

$$M = a(p-q) + 2dpq \qquad [7.2]$$

这既是关于这个性状的群体基因型均值又是它的群体表型均值。

因此，任何一个位点对群体平均值的贡献有两项：$a(p-q)$可归因于纯合体，$2dpq$ 可归因于杂合体。如果无显性（$d = 0$），第二项为 0，则平均值与基因频率成比例：$M = a(1-2q)$。如果存在完全显性（$d = a$），那么平均值与基因频率的平方成比例：$M = a(1-2q^2)$。在没有超显性的情况下，归因于位点值的总范围是 $2a$。也就是说，如果

A_1 在群体中被固定（$p=1$），群体平均值将是 a，如果 A_2 被固定（$q=1$），群体平均值将是 $-a$。可是，如果位点表现超显性，一个没有被固定的群体平均值可能超出这个范围（表 7-2）。

表 7-2 两个等位基因的频率和基因型值

基因型	频率	基因型值	频率×基因型值
A_1A_1	p^2	$+a$	p^2a
A_1A_2	$2pq$	d	$2pqd$
A_2A_2	q^2	$-a$	$-q^2a$
		和=	$a(p-q)+2dpq$

基因型值 a 和 d 是与两个纯合体平均值的离差，如图 7-1 所示。它表明等式［7.2］表示的群体平均值是与两纯合体中点值的离差，两纯合体中点值是尺度的原点或零点。如果平均值要被表示为与某个其他值的离差，那么必须加上或减去一个适当的常数。例如，人们可能想把平均值表示为与较低纯合体值的离差，这将需要加上 a，在一些简化之后，平均值会变成 $M=2p(a+dq)$。或者，被表示为与较高纯合体的离差，它将是 $M=2q(-a+dp)$。

例 7.2

让我们再次取例 7.1 中描述的小鼠的侏儒基因，看看当基因以两个特定的频率出现时，这个基因对群体平均值有什么影响。首先，总范围是从 6g 到 14g：完全由侏儒纯合体组成的群体具有大小为 6g 的平均值，来自完全缺乏侏儒基因的群体具有大小为 14g 的平均值。（这些值专指进行观察时MacArthur's小品系的数值。）现在假定基因以 0.1 的频率出现，因此在随机交配下纯合体将以 1%的频率出现。要代入到等式［7.2］中的数值是 $p=0.9$，$q=0.1$，$a=4$g，$d=2$g，正如例 7.1 所显示的。因此根据等式［7.2］，群体平均值是：$M=4\times0.8+2\times0.18=3.56$。可是，这种平均值是从两个纯合体的中点值（即原点）度量的，其大小为 10g。因此群体平均值的实际值是 13.56g。下一步假定基因以 0.4 的频率出现。以同样的方式代入，我们发现 $M=1.76$，时它必须加上原点值 10g，给出值为 11.76g。这些数字大致确证，可由携带该基因的品系记录中得到。当基因以大约 0.4 的频率出现时，平均体重大约为 12g。两个世代过后，当侏儒基因被有意淘汰时，平均体重上升到大约 14g。

现在我们必须把几个位点上的基因贡献加起来，找出它们对平均值的联合影响。这引起了不同位点上的基因如何结合起来对性状产生联合影响的问题。暂时我们假定结合是通过加性进行的，这意味着涉及几个位点的一个基因型值是归属于各单个位点之值的总和。例如，如果 A_1A_1 的基因型值是 a_A，B_1B_1 的为 a_B，那么 $A_1A_1B_1B_1$ 的基因型值是 a_A+a_B。非加性结合的结果将在本章最后解释。于是，对于加性结合，由几个位点的联合影响导致的群体平均值是各个单个位点中每一个的贡献总和，这样：

$$M=\sum a(p-q)+2\sum dpq \qquad [7.3]$$

这既是基因型均值又是表型均值。现在，在没有超显性的情况下，总范围是$2\sum a$。如果令数值增加的所有等位基因都被固定，平均值将是$+\sum a$，如果令数值减少的所有等位基因都被固定，它将是$-\sum a$。这些是群体中潜在变异范围的理论极限。等式［7.3］中的平均值被度量的起点是总范围的中点。这也等于所有位点各自的两纯合体中点的平均值。

例 7.3

作为加性结合的两个位点的一个例子，我们将涉及小鼠的两个颜色基因，Russell（1949）已描述了它们对色素颗粒数目的影响。这是一个数量性状，它反映了皮毛中色素沉着的强度。这两个基因是“褐色”（b）和“极端稀淡”（c^e），白化体系列的一个等位基因。在野生型纯合体中，两个单个的突变纯合体中，以及双突变纯合体中，对每单位容积毛发中的黑色素颗粒数目进行了度量。我们将假定两种野生型等位基因是完全显性的，因此只需考察4种基因型。4种基因型的颗粒平均数目显示在表7-3。

表 7-3 加性结合的两个位点对色素颗粒的影响

	B-	bb	$2a_B$
C-	95	90	5
c^ec^e	38	34	4
$2a_c$	57	56	

每一行和每一列中两个数字间的差异度量了纯合体差异，或者在图7-1指定数值的尺度上它是$2a$。除了1个单位的轻微差异之外，这些差异与另一个位点的基因型无关。换句话说，B-和bb之间的数值差异，在C-基因型之间与在c^ec^e基因型之间是一样的；类似地是，C-和c^ec^e之间的差异在B-中与在bb中也是一样的。因此，这两个位点可加性地结合起来，一个复合基因型的数值可正确地从已知的单个基因型数值中来预测。例如，bb基因型比野生型少5个单位，c^ec^e基因型比野生型少57个单位；因此，bbc^ec^e应当比野生型的数值95少62个单位，即为33，它与观察值34几乎完全相同。

平均效应

为了推导一个群体与它家系结构相联系的特性，我们必须处理从双亲到后代的数值传递，仅仅借助基因型值是不可能做到这一点的，因为双亲仅传递它们的基因而不是它们的基因型给下一世代，在每一世代中基因型重新创造。因此需要数值的一种新度量，它将涉及到基因而不是基因型。这使得我们能够给个体指定“育种值”，它是与由个体携带并传递给其后代的基因相联系的一个数值。与基因相联系的、而与基因型分开的新数值被称为平均效应。平均效应取决于以前定义的基因型值a和d，还取决于基因频率。因此平均效应是群体和有关基因的特性。平均效应的概念不容易领会，但对理解数量性状的遗传，它是十分重要的。有几种方法可用来定义平均效应。在随机交配下，它们全是等价的，否则不等价（见 Falconer，1985）；这里我们只关注随机繁殖群体。一种定义是这样的：一个特定基因（等位基因）的平均效应是一些个体与群体平均值的平均离差，这些个体从一个亲本接受了该基因，而从另一个亲本接受的基因却是随机的来自于群体。这可用另一种方式

来表述。让携带 A_1 的大量配子全部随机地与来自群体的配子结合；然后如此产生的基因型的平均值和群体平均值的离差数量，就是 A_1 基因的平均效应。

表 7-4 平均效应的推导

配子类型	产生的基因型频率和数值			产生的基因型的平均值	要减去的群体平均值	基因的平均效应
	A_1A_1	A_1A_2	A_2A_2			
	a	d	$-a$			
A_1	p	q		$pa+qd$	$-[a(p-q)+2dpq]$	$q[a+d(q-p)]$
A_2		p	q	$-qa+pd$	$-[a(p-q)+2dpq]$	$-p[a+d(q-p)]$

让我们看看，用来表达群体平均值的基因型值 a 和 d 是如何与平均效应关联起来的。这将有助于使概念更加清楚。推导展示在表 7-4。考虑频率分别为 p 和 q 的两个等位基因 A_1 和 A_2 的一个位点，首先处理基因 A_1 的平均效应，我们将使用符号 α_1 来表示它。如果携带 A_1 的配子随机地与来自群体的配子结合，那么产生的基因型的比例将是 A_1A_1 为 p，A_1A_2 为 q。A_1A_1 的基因型值是 $+a$，A_1A_2 的基因型值是 d，考虑它们发生的比例，这些基因型的平均值是 $pa+qd$。这个平均值与群体平均值之间的差异就是基因 A_1 的平均效应。从等式［7.2］中获取群体平均值的数值，得到：

$$\begin{aligned}\alpha_1 &= pa+qd-[a(p-q)+2dpq]\\ &= q[a+d(q-p)]\end{aligned} \qquad [7.4a]$$

类似地，基因 A_2 的平均效应是：

$$\alpha_2 = -p[a+d(q-p)] \qquad [7.4b]$$

当有两个以上的等位基因时，每一个等位基因的平均效应可以用类似的方式表达。为什么平均效应取决于基因频率的理由，可从词语“随机抽取”的定义中看出，因为一个随机样本的含量取决于群体的基因频率。

当考察一个位点上只有两个等位基因时，采用基因替代的平均效应来表示它们的平均效应更方便。这简单地就是两个等位基因平均效应之间的差异，但采用下面的方法可更清楚地理解它的意义。假定我们可以将随机选择的 A_2 基因改变成 A_1 基因，好像通过定向突变一样，然后我们能注意到由此引起的数值的变化；这样产生的平均变化将是基因替代的平均效应。当 A_2 基因被随机选取时，将在 A_1A_2 基因型中发现比例为 p（p 是 A_1 的基因频率），在 A_2A_2 基因型中发现比例为 q。将 A_1A_2 改变成 A_1A_1 将引起值从 d 变化到 $+a$，因此效应将是 $(a-d)$。将 A_2A_2 改变成 A_1A_2 将引起值从 $-a$ 变化到 d，效应将是 $(d+a)$。因此平均变化是 $p(a-d)+q(d+a)$，经过重新整理，它变成 $a+d(q-p)$。这样，基因替代的平均效应（记为 α，没有下标）是：

$$\alpha = a+d(q-p) \qquad [7.5]$$

根据等式［7.4］，很容易看出：

$$\alpha = \alpha_1 - \alpha_2$$

当采用基因替代的平均效应形式表达时，两个等位基因的平均效应是：

$$\left.\begin{aligned}\alpha_1 &= q\ \alpha\\ \alpha_2 &= -p\ \alpha\end{aligned}\right\} \qquad [7.6]$$

基因型的育种值可方便地用这些项来表达，这将在下一部分看到。基因替代的平均效应的另一种定义将在以后给出，图 7-2 中有图解表示。

例 7.4

再次考察侏儒基因以及它对体重的影响，在例 7.1 中发现其效应是 $a = 4$g 和 $d = 2$g。如果 pg 基因的频率是 $q = 0.1$，根据等式 [7.5]，基因替代的平均效应是 $\alpha = 4 + 2(0.1 - 0.9) = 2.4$g。如果频率是 $q = 0.4$，则基因替代的平均效应是 $\alpha = 4 + 2(0.4 - 0.6) = 3.6$g。根据等式 [7.6]，各个基因的平均效应如表 7-5 所示。

表 7-5 侏儒基因的平均效应

	$q = 0.1$	$q = 0.4$
+ 的平均效应：$\alpha_1 =$	+0.24	+1.44
pg 的平均效应：$\alpha_2 =$	-2.16	-2.16
$\alpha = \alpha_1 - \alpha_2 =$	2.40	3.60

因此，当基因频率较大时，基因替代的平均效应 α 就越大。在两种基因频率上 pg 的平均效应相同只是一种巧合。

育种值

平均效应概念的有用性从已经注意到的事实中产生，即双亲传递它们的基因而不是它们的基因型给它们的后代。因此正是双亲基因的平均效应决定了其后代的平均基因型值。根据其后代平均值来判断一个个体的值被称为个体的育种值。因此不同于平均效应，育种值是可以度量的。如果一个个体与从群体中随机抽取的大量个体交配，那么它的育种值是后代与群体平均值的平均离差的 2 倍。离差必须加倍，因为正被谈论的亲本仅为后代提供一半的基因，另一半基因随机地来自群体。育种值可以用绝对单位表示，但正如上面所定义的那样，通常用与群体平均值的离差形式表示更方便。正像平均效应是基因和群体的一种特性一样，育种值也是个体和它的配偶所属群体的一种特性。没有指定它要交配的群体，人们不能谈及个体的育种值。

表 7-6 基因型与个体育种值

基因型	育种值
A_1A_1	$2\ \alpha_1 = 2q\ \alpha$
A_1A_2	$\alpha_1 + \alpha_2 = (q - p)\ \alpha$
A_2A_2	$2\ \alpha_2 = -2p\ \alpha$

采用平均效应的形式来定义，个体的育种值等于它携带的基因平均效应的总和，即对每一个位点上一对等位基因以及对所有位点求和。因此，对具有两个等位基因的一单个位点，基因型的育种值如表 7-6 所示。

例 7.5

让我们借助小鼠的侏儒基因来说明育种值。+ 和 pg 基因的平均效应给出在上一个例子中。从这些数字中，我们可以找出上面所解释的 3 种基因型的育种值。下面给出的这些

育种值是与群体平均值的离差（表7-7）。基因频率为0.1和0.4的群体平均值可在例7.2中找到，并被再次显示在下面标有 M 的栏目中。

（侏儒纯合体的育种值只是假设的，因为事实上侏儒纯合体几乎全是不育的，但这种复杂情况在本文中可被忽略。）

表7-7　小鼠侏儒基因的育种值

	M	育种值		
		+ +	+ pg	pg pg
$q=0.1$	13.56	+0.48	-1.92	-4.32
$q=0.4$	11.76	+2.88	-0.72	-4.32

扩展到具有两个以上等位基因的一个位点是容易进行的，任一基因型的育种值都是现存的两个等位基因平均效应的总和。如果要考虑所有的位点，一特定基因型的育种值就是归属于每一个单独位点的育种值的总和。如果存在基因型值的非加性结合，则可出现一种稍微复杂的情况。我们给出了育种值的两种定义，用后代的度量值表示的实用上的定义以及用平均效应表示的理论上的定义。非加性结合致使这两种定义不完全等价。这一点将在第9章更全面地解释。

对育种值定义的考虑将表明，在处于哈迪—温伯格平衡的群体中，平均育种值一定为0；或如果育种值用绝对单位来表示，则平均育种值一定等于平均基因型值，也一定等于表型均值。这可以从上面列出的育种值中得到证实。将育种值乘以每一个基因型的频率并求和给出平均育种值（用与群体平均值的离差表达）为：

$$2p^2q\ \ \alpha+2pq(q-p)\alpha-2q^2p\ \ \alpha=2pq\ \ \alpha(p+q-p-q)=0$$

育种值有时被称为“加性基因型”，育种值的变异可归因于基因的“加性效应”。虽然我们将不使用这些术语，但我们在下一章要讨论育种值变异过程时，将遵循习惯使用“加性”这一术语，并且我们将使用符号 A 来标明个体的育种值。

因为育种值表达了从亲本传递到后代的值，所以它表明任何个体预期的育种值是其双亲育种值的平均值；根据育种值的定义，这也是个体的预期表型值。根据后代接受到每一个亲本的两个等位基因的不同，相同亲本的不同后代其育种值将有所不同；期望值简单地就是相同亲本的巨大数量后代的平均值。因此，从亲本到后代传递的数值可表示为：

$$\overline{P_o}=\overline{A_o}=\frac{1}{2}(A_s+A_d) \qquad [7.7]$$

其中，下标 o、s 和 d 分别指后代、公畜和母畜。

显性离差

我们已把作为个体基因型值的一个组分—育种值分离出来了。现在让我们考察什么组成了剩余部分。当只考察一单个位点时，一特定基因型的基因型值 G 和育种值 A 之间的差异被称为显性离差 D，因此：

$$G=A+D \qquad [7.8]$$

显性离差由一个位点上等位基因之间的显性特性而产生，因为在无显性的情况下，育种值

和基因型值相符合。从统计学的观点来看，显性离差是等位基因间的互作或位点内互作。它们代表将基因以成对形式加在一起组成基因型的效应；被单个抽取的两个基因的效应不能计入这种效应。因为基因的平均效应和基因型的育种值都取决于群体的基因频率，因此显性离差也取决于基因频率。所以它们是群体的部分特性，不仅仅是显性度的度量。

例 7.6

继续侏儒基因这个例子，现在我们可以列出基因型值和育种值，因此根据等式 [7.8] 可获得 3 种基因型的显性离差。这些值，现在全部被表示为与群体平均值 M 的离差，给在表 7-8 中。

表 7-8 侏儒基因型的基因型值、育种值和显性离差（表示为与群体平均值的离差）

	$q=0.1$; $M=13.56$			$q=0.4$; $M=11.76$		
	+ +	+ pg	pg pg	+ +	+ pg	pg pg
频率	0.81	0.18	0.01	0.36	0.48	0.16
基因型值 C	+0.44	−1.56	−7.56	+2.24	+0.24	−5.76
育种值 A	+0.48	−1.92	−4.32	+2.88	−0.72	−4.32
显性离差 D	−0.04	+0.36	−3.24	−0.64	+0.96	−1.44

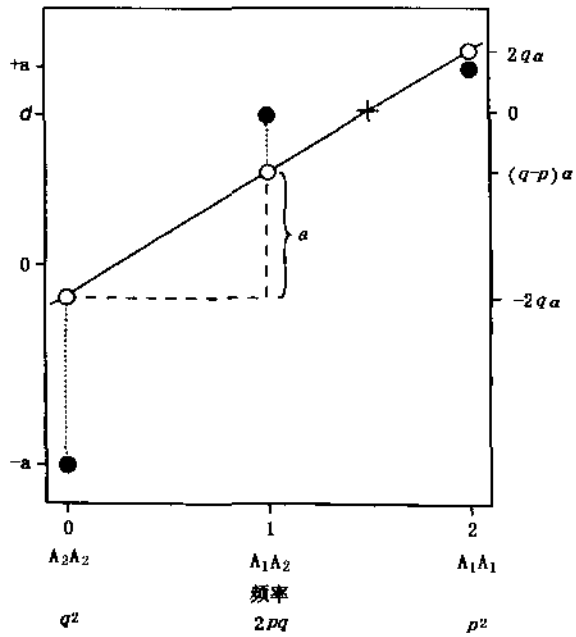


图 7-2 正如正文中解释的，图示为具有频率为 p 和 q 的两个等位基因 A_1 和 A_2 的一个位点上的基因型的基因型值（实心点）和育种值（空心点）

水平尺度：基因型中 A_1 基因数目。数值的垂直尺度：左边，像图 7-1 一样指定的任意值；右边，与群体平均值的离差。该图以下列数值为尺度而绘制：$d=\frac{3}{4}a$，$q=\frac{1}{4}$。

基因型值、育种值和显性离差之间的关系可以用图解来说明，像图 7-2 那样，用这种方式可能更容易理解显性离差的意义。在图中，基因型值（实心点）以基因型中 A_1 基因的数目为背景而绘制。通过最小二乘法对这些点拟合了一条直的回归线，每一点用它代表的基因型频率来加权。这条直线的位置给出了每一种基因型的育种值，用空心点表示。育种值和基因型值之间的差异是显性离差，用垂直点线表示。+字形表示群体平均值。基因替代的平均效应 α 由 A_2A_2 和 A_1A_2 之间或 A_1A_2 和 A_1A_1 之间的育种值差异给出，如图所示。基因替代的平均效应的最初定义就是由 Fisher（1918，1941）采用基因型值对基因数目这种线性回归的形式给出的。

表 7-9 在一个两个等位基因系统中的基因型值，用与群体平均值的离差为度量

	基因型		
	A_1A_1	A_1A_2	A_2A_2
频率	p^2	$2pq$	q^2
指定值	a	d	$-a$
与群体平均值的离差：			
基因型值	$\begin{cases}2q(a-pd)\\2q(a-qd)\end{cases}$	$a(q-p)+d(1-2pq)$ $(q-p)a+2pqd$	$-2p(a+qd)$ $-2p(\alpha+pd)$
育种值	$2q\alpha$	$(q-p)\alpha$	$-2p\alpha$
显性离差	$-2q^2d$	$2pqd$	$-2p^2d$

注：群体平均值：$M=a(p-q)+2dpq$

基因替代的平均效应：$\alpha=a+d(q-p)$。

通过从基因型值中减去育种值，显性离差可以采用任意指定的基因型值 a 和 d 来表示，如表 7-9 所示。基因型值首先必须被转换成与群体平均值的离差，因为育种值已采用这种方法表示了。这样转换过的基因型值以两种形式给出：以 a 的形式和以 α 的形式。让我们考察基因型 A_1A_1 来显示这些是如何获得的，通过减去育种值显性离差又是如何获得的。A_1A_1 的任意指定的基因型值是 $+a$，群体平均值是 $a(p-q)+2dpq$。按与群体平均值的离差表示，因此，基因型值是：

$$a-[a(p-q)+2dpq]=a(1-p+q)-2dpq=2qa-2dpq=2q(a-dp)$$

通过代入 $a=\alpha-d(q-p)$（来自等式 [7.5]），这可以采用平均效应 α 的形式来表示，于是基因型值变成 $2q(\alpha-qd)$。减去育种值 $2q\alpha$，给出显性离差为 $-2q^2d$。通过类似的推导，A_1A_2 的显性离差是 $2pqd$，A_2A_2 的显性离差是 $-2p^2d$。因此所有的显性离差都是 d 的函数。如果无显性，d 是 0，显性离差也全部为 0。因此在无显性的情况下，育种值和基因型值是相同的。不表现显性（$d=0$）的基因有时被称为“加性基因”，或说成是“以加性方式起作用”。

因为平均育种值和平均基因型值是相等的，它表明平均显性离差为 0。这通过将显性离差乘以每一种基因型的频率并求和可以得到证实。这样平均显性离差是：

$$-2p^2q^2d+4p^2q^2d-2p^2q^2d=0$$

互作离差

当只考察一个单个位点时，基因型值仅由育种值和显性离差组成。但当基因型涉及一

个以上位点时，基因型值可能包含由非加性结合引起的一种附加的离差。假定 G_A 表示归因于一个位点的个体的基因型值，G_B 表示归因于第二个位点的基因型值，G 表示归因于两个位点在一起的总基因型值。于是：

$$G = G_A + G_B + I_{AB} \qquad [7.9]$$

其中，I_{AB}是与这些基因型值的加性结合的离差。在本章之初，处理群体平均值的过程中，我们假定对基因型的所有结合 I 都为0。如果对不同位点基因的任何结合 I 不是0，则那些基因被说成“互作”或表现出“上位”，术语上位在数量遗传学中比在孟德尔遗传学中被赋予了更广泛的意义。离差 I 被称为*互作离差*或*上位离差*。如果互作离差是0，有关的基因被说成在位点之间“以加性方式起作用”。因此“加性作用”可能指两种不同的事情。涉及一个位点上的基因，它意味着无显性，涉及不同位点上的基因，它意味着无上位。

位点可能成对地或以3个或更多的数目互作，互作可能有几个不同的类型，就像主效基因的行为所显示的一样。可是，我们不必关注互作的复杂属性，因为在总基因型值中，所有类型的互作放在一起作为单一互作离差来处理。因此，对于放在一起的所有位点，我们可以书写：

$$G = A + D + I \qquad [7.10]$$

其中，A 是归属于单独位点的育种值的总和，D 是显性离差的总和。

当数值用与群体平均值的离差表示时，一个群体中所有基因型的平均互作离差是0。这是肯定的，这一点可从等式［7.10］中看出，记住 G、A、D 的平均值全为0。互作离差不仅是互相作用的基因型的一种特性，而且也取决于群体中基因型的频率，因此也取决于基因频率。

例 7.7

作为两个位点非加性结合的一个例子，我们将选取在例7.3中被用来说明加性结合的小鼠中相同的两个颜色基因；但这次我们涉及它们对色素颗粒大小而不是对它们数目的影响（Russell，1949）。在4种基因型中颗粒的平均大小（直径以 μ 为单位）显示在表7-10中。这次差异与另一种基因型有关：例如，c^e 基因对B－基因型有相当大的影响，但对bb基因型一点也没有影响。因此这两个位点具有上位互作，不是以加性方式结合起来的。在任一特定群体中4种基因型的互作离差将取决于两个位点上的基因频率。

表7-10 四种基因型色素颗粒大小平均值

	B－	bb	差异
C－	1.44	0.77	0.67
c^ec^e	0.94	0.77	0.17
差异	0.50	0.00	

习 题

7.1 红细胞酸性磷酸酶的3种等位变种存在于英国群体的一个样本中。下表给出了样本中具有频率的基因型以及每一种基因型的平均酶活性（CC个体没有发现）。这个群体中平均酶活性是多大？

基因型	频率（%）	酶活性
AA	9.6	122
AB	48.3	154
BB	34.3	188
AC	2.8	184
BC	5.0	212

数据来自 Spencer，N.，*et al*.（1964）*Nature*，**201**，299～300。

[解答序号：6]

7.2 对习题 7.1 中给出的红细胞酸性磷酸酶的酶活性，在没有 C 等位基因和 A 具有下列基因频率：（1）0.2，（2）0.5，（3）0.8 时，计算群体中的平均酶活性。

[解答序号：16]

7.3 如果有一个对于一个数量性状为超显性的位点，具有下面给出的基因型值，那么多大的基因频率将使一个随机交配群体产生它的最大平均值，这个平均值是多大？

A_1A_1	A_1A_2	A_2A_2
110	150	90

[解答序号：26]

7.4 例 7.3 描述了两个颜色基因对小鼠毛发中色素颗粒数目的影响。当褐色基因 b 纯合时，将颗粒数从 95 减少到 90，极端稀淡基因 c^e 纯合时将颗粒数从 95 减少到 38。结合的基因效应是加性的。假定两个基因是完全隐性，找出 b 基因频率为 0.5 和 c^e 基因频率为 0.2 的一个群体中平均的色素颗粒是多少。

[解答序号：36]

7.5 在习题 7.2 指定的群体中，两个等位基因的平均效应和基因替代的平均效应各是多大？

[解答序号：46]

7.6 在习题 7.2 指定的 3 个群体中，3 种基因型的育种值是多大？

[解答序号：56]

7.7 当平均值处于它的最大值时，找出习题 7.3 指定的群体中 3 种基因型的育种值和显性离差。

[解答序号：66]

7.8 在习题 7.4 指定的群体中，计算基因型 bbc^cc^c 的育种值和显性离差。同时以与群体平均值的离差以及以颗粒数的绝对单位给出育种值。

[解答序号：76]

7.9 习题 7.4 涉及两个颜色基因对小鼠被毛中色素颗粒数目的影响。例 7.7 描述了相同的两个基因对色素颗粒大小的影响，在这种情形中，效应不是加性的。在 bb 纯合子频率为 0.4 和 c^cc^c 纯合子频率为 0.2 的一个群体中，求解出基因型的互作离差。因为两个基因都被假定为完全隐性的，在每一种情形中显性纯合体和杂合体具有相同的值，可以作为一种基因型来处理，因此，有 4 种基因型，需要找出它们的互作离差。如果例 7.7 给出的 4 种基因型值首先被转换成与群体平均值的离差，这对后面的习题将是有用的；这还可以使条理更清楚。

[解答序号：86]

第 8 章　方　　差

数量性状的遗传学以其变异的研究为中心，因为正是采用了变异的术语才使主要的遗传学问题得以系统地阐述。变异研究的基本思想是将其分剖为归属于不同原因的组分。这些组分的相对大小决定了群体的遗传特性，尤其是亲属间的相似程度。在这一章中，我们将考察这些组分的属性以及遗传组分是如何取决于基因频率的。然后，在下一章，我们将阐明亲属间相似的程度如何由组分的大小所决定。

方差组分

变异的大小可度量并以方差表示：当以偏离群体平均值的离差表示数值时，简单地说方差就是该值平方的平均值。方差分剖成的组分与上一章描述的数值的组分相同；因此，举例来说，基因型方差就是基因型值的方差，环境方差就是环境离差的方差。总方差是表型方差，或表型值的方差，也是各个组分的总和。方差组分以及它们所度量的方差所隶属的数值列于表 8-1。

表 8-1　方差组分

方差组分	符号	被度量方差所隶属的数值
表型	V_P	表型值
基因型	V_G	基因型值
加性	V_A	育种值
显性	V_D	显性离差
互作	V_I	互作离差
环境	V_E	环境离差

于是，有一定限制条件的总方差是各个组分的总和，这样：

$$V_P = V_G + V_E \qquad [8.1a]$$

$$= V_A + V_D + V_I + V_E \qquad [8.1b]$$

限制条件是，首先，基因型值和环境离差可能相关，在这种情况下，V_P 将增加 G 与 E 协方差的两倍；其次，基因型和环境之间可能存在互作，在这种情况下，将有一种归因于互作的额外方差组分。这两种复杂情况将在本章后面论及；同时假定它们不适用。

作为总方差部分的组分

将方差分剖成它的组分有利于我们估计表型的各种决定因子的相对重要性，尤其是遗传对环境，或本质与环境因素。相对重要性的问题只有当它用归因于不同变异来源的方差形式来表达时才能被回答。一种变异来源的相对重要性就是由这种来源引起的方差，占总

表型方差的一种比例。在确定表型值中遗传的相对重要性被称为性状的遗传力。可是，根据它们指的是基因型值还是育种值，遗传和遗传力有两种明显不同的含义。在由基因型决定的意义上或从双亲传递给后代的意义上一个性状均是遗传的，然而性状的遗传程度，就这两种意义而言，却可能是不同的。比率 V_G/V_P 表示个体的表型被基因型所决定的程度。这被称为广义遗传力，或遗传决定的程度。比率 V_A/V_P 表示表型由双亲传递下来的基因所决定的程度。这被称为狭义遗传力，或简单地称为遗传力。总的来说这表明，术语“遗传力”将被限定指狭义遗传力 V_A/V_P。遗传力 V_A/V_P 决定亲属间相似的程度，因此它在育种方案中是最重要的。与实际重要性相比，由 V_G/V_P 决定的遗传程度具有较多的理论重要性。它可采用下面的方法来估计。

遗传决定程度 V_G/V_P 的估计

基因型方差 V_G 的估计在理论上是简单的，虽然在实践中不那么容易。方差的基因型组分 V_G和环境组分 V_E 都不能从一单个群体的观测中被直接估计，但在特定条件下，它们可由实验群体估计出来。如果一种或其他组分能被完全排除，则剩余的表型方差将提供剩余组分的一种估计值。环境方差不能被消除，因为根据定义它包括所有的非遗传方差，这种方差的大部分超出了实验控制。可是，基因型方差的排除，是可以通过实验达到的。从一个高度近交系或两个这样的系之间杂交的 F_1 代，或由一单个个体繁殖的一个克隆中可以获得相同基因型的个体。(人和牛的同卵双生子也提供相同基因型的个体，但由于在第10章中要讨论的理由，在分剖方差中它们的运用是很有限的。) 如果一组这样的个体在正常范围的环境条件下饲养，那么，它们的表型方差提供了环境方差 V_E 的估计值。于是，从一个遗传上的混合群体的表型方差中减去这个方差就给出这个群体基因型方差的估计值。这种估计在下面的例子中进行说明。

例 8.1

在黑腹果蝇中已对几个性状进行了将表型方差分剖成它的基因型和环境组分。结果稍后在表8-4中给出，但为了显示如何进行分剖，这里我们可以用更多的细节来描述一个性状的结果。这个性状是胸长（以 1/100mm 为单位），它被认为是身体大小的一种度量。分别在一个遗传上混合的群体（即随机繁殖群体）中以及在由高度近交系3次杂交的 F_1 代组成的遗传上一致的群体中测量了表型方差。第一个群体估计了基因型方差和环境方差的总和，第二个群体仅仅估计了环境方差，正如表8-2中所示。因此，借助减法，可以获得基因型方差的估计值。这表明，遗传上混合群体中胸长总变异的49%归因于个体之间的遗传差异，51%归因于非遗传差异（数据来自 F.W.Robertson，1957*b*）。

表 8-2 黑腹果蝇胸长性状的方差组分分剖

群体	组分	观察的方差
混和	$V_G + V_E$	0.366
一致	V_E	0.186
差异	V_G	0.180
	$V_G/V_P =$	0.180/0.366 = 49%

以上面描述的方式进行基因型方差的估计并不像看起来那么容易进行。它依赖于所有基因型中的环境方差都相同这一假设，这当然不总是正确的。在一个近交系或杂种中测量

的环境方差，是这个特定的基因型所显示的方差，其他的基因型可能对环境影响更敏感或不敏感，因此可能表现出或更大的或更小的环境方差。因此，混合群体的环境方差可能与在基因型一致的群体中所测量的环境方差不相同。而且，已经发现一些性状在近交种、纯合个体之间比在杂交种、杂合个体之间更富于变化，纯合体对环境差异更敏感。对环境的不同敏感性是基因型—环境互作的一个方面，而这种互作将在本章后而更全而地讨论。因为可能有环境敏感性的差异，所以从几种不同的一致群体，近交系和它们的杂交组合中来估计环境方差是人们所理想的。联合单独的估计值以得到最可靠的平均估计值的方法，见 Wright（1968，p.382）。

可借助能无性繁殖的植物来克服因运用不典型的基因型所引起的困难。因此，可能有大量的基因型，每一个基因型可由很多无性产生的个体来代表。如果就环境而言个体成长为随机化的基因型，那么方差分析可提供无性系之间方差组分的估计值。无性系之间的方差主要由基因型差异引起，可被认为是 V_G 的估计值。但在 V_G 中可能包括有环境效应。也就是说，个体之间环境差异的一部分可能传递给它们所有的无性系后代。由于这个原因，用这种方法估计的比率 V_G/V_P 倾向于是一个过高估计。严格地讲，它应当被称为“无性系重复力”；重复力的含义将在本章后而解释。

方差的遗传组分

分剖成基因型方差和环境方差并没有使我们更深入地去了解群体的遗传特性，尤其是它并没有揭示亲属间相似的原因。基因型方差必须根据基因型值再细分成育种值、显性离差和互作离差。这样有：

数值	G =	A +	D +	I	
方差组分	V_G =	V_A +	V_D +	V_I	[8.2]
	（基因型）	（加性）	（显性）	（互作）	

加性方差 它是育种值的方差，是重要的组分，正如已提到的那样，因为它是亲属间相似性的主要原因，因而是群体可测遗传特性以及群体对选择起反应的主要决定因子。而且，它是很容易从群体获得的观测值中迅速估计出来的唯一组分。因此，在实践中，重要的分剖是将加性遗传方差和其余方差分开，其余的是非加性遗传方差和环境方差。这种分剖产生了比率 V_A/V_P，它是性状的遗传力。

加性方差的估计依赖于对亲属间相似程度的观测，这将在我们讨论过亲属间相似的原因以后予以描述。这里我们即刻关心的是看看方差的遗传组分如何受基因频率的影响。要做到这一点，我们必须用基因频率和指定的基因型值 a 和 d 来表示方差。首先我们将考察具有两个等位基因的一单个位点，这样暂时排除了互作方差。

加性和显性方差

获得育种值方差和显性离差方差的表达式所需要的信息给在上一章的表 7-9 中。该表给出的 3 种基因型的育种值和显性离差是用与群体平均值的离差表示的。要记住育种值和显性离差两者的平均值都为 0。因此，不必对假定平均值进行校正，方差简单地就是平方值的平均值。这样，通过对表中的值进行平方、乘上有关基因型的频率并对 3 种基因型求和，可获得方差。获得的加性方差即育种值方差如下：

$$\begin{aligned} V_A &= 4p^2q^2\alpha^2 + 2pq(q-p)^2\alpha^2 + 4p^2q^2\alpha^2 \\ &= 2pq\ \alpha^2(2pq + q^2 - 2pq + p^2 + 2pq) \\ &= 2pq\ \alpha^2(p^2 + 2pq + q^2) \\ &= 2pq\ \alpha^2 \qquad [8.3a] \\ &= 2pq[a + d(q-p)]^2 \qquad [8.3b] \end{aligned}$$

像表 7-9 中一样，这里 q 是降低数值的等位基因频率。

类似地，显性离差的方差是：

$$\begin{aligned} V_D &= d^2(4q^4p^2 + 8p^3q^3 + 4p^4q^2) \\ &= 4p^2q^2d^2(q^2 + 2pq + p^2) \\ &= (2pqd)^2 \qquad [8.4] \end{aligned}$$

如果在所考察的位点上无显性（$d=0$），那么加性方差的表达式简化成：

$$V_A = 2pq\alpha^2 \qquad [8.5]$$

如果存在完全显性（$d=a$），则加性方差变为：

$$V_A = 8pq^3\alpha^2 \qquad [8.6]$$

其中 q 是隐性等位基因的频率。

对任何程度的显性，当它们位于由两个高度近交系的杂交获得的群体中时，如果所有分离基因的频率都是 1/2（$p=q=0.5$），那么加性和显性方差两者的表达式变得简单得多。于是，等式 [8.3b] 和 [8.4] 简化成：

$$\left.\begin{aligned} V_A &= \frac{1}{2}\alpha^2 \\ V_D &= \frac{1}{4}d^2 \end{aligned}\right\} \qquad [8.7]$$

对这类群体分析的全面考察，参见 Mather 和 Jinks（1977，1982）。

总遗传方差

由一个位点产生的总遗传方差 V_G 可以采用与上面对 V_A 和 V_D 进行的相同方法直接从表 7-9 中计算得到。但这样做需要冗长的代数约简，从上面计算的 V_A 和 V_D 值中获得 V_G 更简单。因为 $G=A+D$，G 的方差由 $V_G = V_A + V_D + 2\mathrm{cov}_{AD}$给出，其中 cov_{AD}是育种值和显性离差的协方差。这个协方差可以显示为 0，如下所示，三种基因型的育种值、显性离差和频率给在表 7-9 中。将育种值乘以显性离差再乘以频率，对 3 种基因型求和，可给出协方差为：

$$-4p^2q^3\alpha d + 4p^2q^2(q-p)\ \alpha d + 4p^3q^2\alpha d = 4p^2q^2\alpha d(-q+q-p+p) = 0$$

因此

$$\begin{aligned} V_G &= V_A + V_D \\ &= 2pq[a + d(q-p)]^2 + [2pqd]^2 \qquad [8.8] \end{aligned}$$

例 8.2

为了说明由单个位点引起的方差遗传组分，让我们回到在上一章中用作几个例子的小鼠侏儒基因上来。根据例 7.6 中列成表的数值，我们可以直接计算方差组分。因为数值是

以与群体平均值的离差表示的，所以通过将每一种基因型的频率乘以其值的平方，并对3种基因型求和，就可获得方差。例如，当 $q=0.1$ 时，基因型方差是 $0.81\ (0.44)^2+0.18\ (-1.56)^2+0.01\ (-7.56)^2=1.1664$。加性方差可用同样方法从育种值的方差中获得，显性方差同样可从显性离差的方差中获得。获得的方差如表8-3所示。

表8-3 小鼠侏儒基因位点的加性方差和显性方差

	$q=0.1$	$q=0.4$
基因型方差 V_G	1.1664	7.1424
加性方差 V_A	1.0368	6.2208
显性方差 V_D	0.1296	0.9216

运用上面在等式［8.3］、［8.4］和［8.8］中给出的公式，也可获得方差，在计算上困难还少些。要代入的数值给在例7.1中；即 $a=4$ 和 $d=2$。注意显性方差与加性方差相比是很小的。

基因频率和显性度影响方差遗传组分大小的方式，可从上述等式［8.3］、［8.4］和［8.8］获得的关系图中充分地察知。图8-1的图解显示了由具有两个等位基因的一单个位点引起的基因型、加性和显性方差的大小，它们是以基因频率为背景绘制成图的。显示出3种情况用来说明不同显性度的影响：在图（a）中无显性（$d=0$）；在图（b）中存在完全显性（$d=a$）；在图（c）中存在纯粹的超显性（$a=0$）。在第一种情形下，基因型方差全是加性的，当 $p=q=0.5$ 时，它达到最大。在第二种情形下，当 $p=q=0.5$ 时，显性方差最大，当隐性等位基因的频率 $q=0.75$ 时，加性方差最大，当 $q^2=0.5$，即 $q=0.71$ 时，基因型方差最大。在第三种情形下，显性方差与第二种情形下相同，当 $p=q=0.5$ 时，达到最大值。可是，当 $p=q=0.5$ 时，加性方差为0，加性方差有两个最大值，一个位于 $q=0.15$，另一个位于 $q=0.85$。在这种情形下，基因型方差实际上在很大的基因频率范围内保持稳定不变，虽然它的组分有深刻变化。从这些图中得出的一般结论是当基因具有中间频率时，它们对方差的贡献比具有高或低频率时要大得多；尤其是，低频率的隐性基因贡献极少的方差。

前面对遗传方差的讨论主要是理论性的。在实践中，我们不关心基因频率或基因效应，因为它们是不知道的，只是特殊构建的群体除外。因此，在实践中，我们只关心组分的估计。可是应当注意到，遗传方差的所有组分都取决于基因频率，因此它们的任何估计值仅对它们被估计的群体有效。当有关亲属间相似性的观测数据有效时，我们就能估计加性遗传方差（估计方法是下面两章的主题）。于是总表型方差可分剖成 $V_A:(V_D+V_I+V_E)$，非加性遗传方差包括在环境方差中。如果近交系是可用的，则可估计环境方差，这样表型方差可分剖成 $V_G:V_E$。

如果这两种分剖都做过，那么我们能将加性遗传方差同遗传方差的剩余部分分离开来，因此而进行加性遗传、非加性遗传和环境方差 $V_A:(V_D+V_I):V_E$ 的三重分剖，即将显性和互作组分合在一起作为非加性遗传方差。这种分剖的例子给在表8-4中。

在这里应当提及，术语可能引起对于加性遗传方差概念的一种误解，加性方差的概念

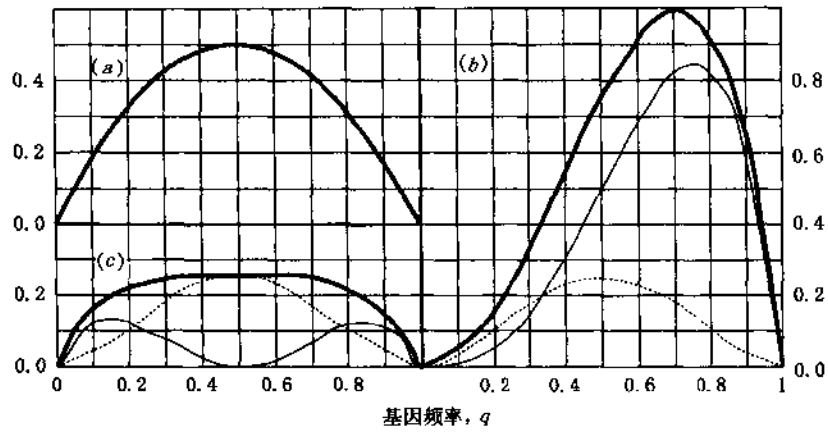


图 8-1　两个等位基因的一单个位点产生的方差遗传组分的大小与基因频率的关系

基因型方差——粗线；加性方差——细线；显性方差——虚线。基因频率 q 是隐性等位基因的频率。显性度是：在（a）中无显性（$d=0$）；在（b）中完全显性（$d=a$）；在（c）中纯粹的超显性（$a=0$）。垂直尺度上的数字，表示方差的数量，在图（a）和（b）中须乘以 a^2，在图（c）中须乘以 d^2。

并不附有加性基因作用的假设，并且加性方差的存在不是任何基因加性作用（即，表明既不是显性又不是上位）的指标。对有关的基因作用方式没有作出假设。加性方差可由具有任何显性度或上位的基因所产生，只有当我们发现所有的基因型方差都是加性时，我们才能得出结论：那些基因既不是显性也不是上位。

对一个位点上存在两个以上等位基因的情况不引用新的原理，虽然它使得位点效应的理论描述复杂化了。Kempthorne（1955a）给出了加性和显性方差的表达式。位点贡献出加性方差和显性方差，前者产生于若干等位基因的平均效应，后者产生于若干显性离差。

表 8-4　黑腹果蝇（*Drosophila melanogaster*）中 4 个性状的方差分剖（组分用总表型方差的百分率表示）

		性　状			
		(1) 刚毛	(2) 胸长	(3) 卵巢	(4) 卵
表型方差	V_P	100	100	100	100
加性遗传方差	V_A	52	43	30	18
非加性遗传方差	V_D+V_I	9	6	35	44
环境方差	V_E	39	51	35	38

性状和数据来源：

(1) 在第 4、第 5 腹节上的刚毛数（Clayton，Morris 和 Robertson，1957；Reeve 和 Robertson，1954）。

(2) 胸长（Robertson，1957b）。

(3) 卵巢大小，即两个卵巢上卵巢管的数目（Robertson，1957a）。

(4) 4 天内（羽化后第 4 到第 8 天）产卵数（Robertson，1957b）。

为了得出群体表达的方差组分，对方差有贡献的所有位点的单独效应必须结合在一起。当一个随机交配群体处于平衡时，由所有位点一起产生的加性方差就是由每一个位点

单独产生的加性方差的总和；类似地，显性方差也是各自贡献的总和。但考察一个以上的位点时，如果存在互作离差，那么会产生另一种方差组分，即互作方差，它是互作离差的方差。

互作方差

如果以前一章描述的方式在不同位点上的基因型表现出上位互作，那么互作会产生方差组分 V_I，它是互作离差的方差。互作方差特性的理论描述依赖于它进一步细分成的组分。首先根据所包含的位点数而被细分为：由两个位点的互作引起两因子互作，由三个位点引起三因子互作，等等。因为由较大数目的位点互作只产生很小的方差，以致于我们可将之忽略不计，所以我们将注意力集中到两因子互作，因为这种互作足以用来说明所包含的原理。互作方差的下一步细分需根据互作包含育种值还是显性离差而定。这样存在 3 种类型的两因子互作。两个育种值之间的互作产生加性 × 加性方差，V_{AA}；一个位点的育种值和另一个位点的显性离差之间的互作产生加性 × 显性方差，V_{AD}；两个显性离差之间的互作产生显性 × 显性方差，V_{DD}。因此互作方差被分剖成如下的组分：

$$V_I = V_{AA} + V_{AD} + V_{DD} + \cdots \qquad [8.9]$$

标明“…”的项是指由两个以上位点之间的互作引起的类似组分。

毫无疑问，控制数量性状的位点之间的互作是经常发生的：例如，它已经在果蝇（*Drosophila*）的很多研究中被证明了（见 Kearsey 和 Kojima，1967）。可是，估计它产生的方差大小是不容易的，对于它作为一种变异来源的相对重要性知道甚少。Barker（1979）综述了实验证据。上位互作的进一步细节参见 Cockerham（1954，1963），Kempthorne（1957），Crow 和 Kimura（1970）。

在通过相对简单的实验，例如这里所例举的，所进行的方差分剖过程中，大部分的互作方差被包括在显性组分中，于是它被称为*非加性遗传方差*。对于互作方差的描述，我们在这里先论述这些，但在下一章我们将看到它是如何对亲属间的相似性作贡献的。

由不平衡引起的方差

在这个段落中必须提及遗传方差的一种额外来源，虽然它仅在以后几章的很少地方与我们有关。当一个群体不处于随机交配下的平衡时，它就产生了。当同时考察两个以上位点上的基因型频率不是依基因频率得来的期望频率时，不平衡就会存在。由于下面的原因，不平衡引入了遗传方差的一种额外来源。为简单起见，仅考察两个位点，它们之间没有上述方式的互作。令 G' 和 G'' 分别表示与每个位点有关的个体基因型值，令 G 表示两个位点联合起来的基因型值，即 $G = G' + G''$。于是，由两个位点共同引起的总基因型方差是：

$$V_G = V_{G'} + V_{G''} + 2\mathrm{cov}_{G'G''} \qquad [8.10]$$

协方差项表示不同个体中两个位点上的基因型值之间的相关。相关可以是正的或负的，因此不平衡可以增加也可以减少方差。当考察两个以上位点时，每一对位点都将有一个协方差项。当不存在不平衡时，所有的协方差项为 0，方差和前面部分描述的相同。

有两种形式的非随机交配可以产生不平衡，它们以不同的方式产生等式［8.10］中的协方差。当亲本不是它们世代中个体的一个随机样本时，第一种非随机交配就发生了。亲

本的选择，构成了这种类型的非随机交配，它是后面几章的主题。第二种形式的非随机交配是选型交配，正如第 1 章所描述的那样。产生的两种类型的协方差代表了基因效应的不同相关。首先，同一配子不同位点上的基因之间存在一种相关。这是配子期不平衡或连锁不平衡，这在第 1 章解释过。第二种类型的协方差代表在配子结成对中基因之间的相关，即一个体从两个亲本中接受的基因之间的相关。单独的非随机交配的第一种形式，即亲本选择后它们再随机交配，产生了第一种类型的协方差，它是由配子期不平衡引起的。第二种形式的非随机交配，即选型交配，产生了两种类型的协方差。如果不平衡的来源停止起作用，那么不是由配子期不平衡引起的协方差会立即消失。不连锁位点的配子期不平衡在每一个随后的世代中减少一半，但对于连锁位点，它持续较长时间。

基因型和环境之间的相关与互作

两种复杂情况的出现与方差分剖成等式［8.1］表示的基因型和环境方差有关。正常地忽略它们，不会严重影响从分剖方差中得出的结论，但重要的是要知道忽略它们的后果是什么。

相 关

在前面对方差组分的考察中，假定环境离差和基因型值是互相独立的；换句话说，基因型值和环境离差之间没有相关。如果较好的环境产生较好的基因型，那么，它们之间有相关的情况就会出现。基因型和环境之间的相关是一个少见的重要而复杂的问题，通常在实验群体中被忽略，在那里环境的随机化是实验设计的主要目标之一。可是，有一些情形中存在这种相关。奶牛的产奶量是一个例子。乳牛业正常实践是根据母牛的产量来饲喂母牛，对较好的表型给予较多的饲料。这引起了表型值和环境离差之间的相关；并且因为基因型值和表型值是相关的，所以基因型值和环境离差之间也存在相关。另一个例子是人的智力。双亲的表型值影响孩子成长的环境；因此，在某种程度上说，智力是可遗传的，这引起了孩子的基因型和环境之间的相关。只有在环境离差和基因型值不相关时等式［8.1a］才是正确的。当相关存在时，表型方差增加基因型值和环境离差协方差的两倍，等式［8.1a］变成：

$$V_P = V_G + V_E + 2\mathrm{cov}_{GE} \qquad [8.11]$$

cov_{GE}可度量的唯一方法是，如果我们像例 8.1 那样直接估计出 V_E，又在固定环境中通过变化的基因型方差也直接估计出 V_G，那么按等式［8.11］从 V_P 中减去直接估计的 V_G 和 V_E 将产生 $2\mathrm{cov}_{GE}$的估计值。可是，只有近交系是有效的并且环境的非随机方面（例如饲喂水平）可以识别，才能做到这一点。在实践中未知的协方差最好被看作是遗传方差的一个部分，因为环境的非随机方面是基因型值的一种后果，因此个体的环境可被认为是其基因型的一部分。这对母牛的产奶量不是不合理的。对人的智力它不那么令人满意，因为环境对孩子的影响不是它们自身基因型的结果，而是双亲基因型的结果。

互 作

已经作出的另一个假设是，环境的一种特殊差异对不同基因型具有相同的影响，它不

总是无可非议的；或者换句话说，我们能够将一种特定环境离差和环境的一种特殊差异联系起来，而不管它起作用的基因型是什么。当情况不是这样时，在统计意义上就说基因型和环境之间存在一种互作。这种互作可能采取的形式有几种。例如，环境的一种特殊差异可能对一些基因型比对另一些基因型具有较大的影响；或者在不同环境下度量时，一系列基因型优势的次序可能会发生变化。也就是说，在环境 X 中基因型 A 可能优于基因型 B，但在环境 Y 中却次于基因型 B。

当基因型和环境之间存在互作时，个体表型值不是简单地像等式［7.1］表示的那样，$P = G + E$，而是还包括了一个互作组分：$P = G + E + I_{GE}$。互作组分产生了一种额外的变异来源，等式［8.11］变成 $V_P = V_G + V_E + 2\mathrm{cov}_{GE} + V_{GE}$。

在例 8.1 中说明的这类实验中，遗传上一致的群体是单一基因型的。它的方差完全是由个体间环境差异引起的，取决于特定基因型对环境差异起反应的方式。因此由互作引起的方差包括在从该基因型的表型方差中估计出的环境方差之中。已经注意到，一些基因型可能比另一些基因型对环境差异更敏感。因此，在某种程度上，环境方差是基因型的一种特性。但变异来源是环境的，而非遗传的。因此，将基因型—环境互作引起的任何方差看作是包括在 V_E 任何估计值中的环境方差的一部分是合乎逻辑的，在实验上也是必要的。

如果一特定群体的个体将在不同条件下饲养，那么基因型—环境互作变得很重要。例如，家畜的一个品种可能被不同的农场主所使用，而这些农场主会不同地对待它：各种植物生长在不同的季节、不同的地方和不同的条件下。这种情形在一个方面不同于迄今我们已考察的情况。不同的农场、季节或地方是“特殊环境”，由位于它们之中的所有个体分享，在处理的属性方面更多。在迄今所考察的情形中，每一个体都有它自己的特定环境，根据它们环境的任何特定方面，例如营养、温度或拥挤状况，不能对个体分组。当个体饲养在特殊环境中时，可更加详细地研究基因型—环境互作。如果基因型可以复制，几种基因型中每一种基因型都有一个以上个体饲养在不同的特殊环境中，那么采用基因型 × 环境的二向分类的方差分析将产生基因型之间方差、特殊环境之间方差、归因于基因型和环境的互作方差的估计值。如果没有互作，那么在一个环境中最好的基因型，在所有环境中都将是最好的。但是如果存在大量的互作，那么对特定环境必须找出特定的基因型。关于特殊环境的品种专化性以后在第 19 章中还要提到，因为从一种不同的观点来讨论它更有用。下一步我们将考察环境敏感性的概念以及它如何被度量的。

环境敏感性 一些基因型—环境互作可归因于不同基因型敏感性的差异。换句话说，一给定的环境差异对一些基因型比对另外一些基因型具有更大的影响（一种基因型的环境敏感性也称为它的“反应规范”）。为了度量环境敏感性，以及看看互作方差有多少可归因于敏感性的差异，将不同基因型饲养在或生长在一定范围的特殊环境中。特殊环境必须限定为对所研究性状的表达或多或少是有利的。对环境定量的唯一方法是用所有基因型的平均表现。换句话说，一个环境的度量是在该环境中所有基因型的平均值。这将被称为环境值。每一种基因型在每一种特殊环境中都有它自己的平均值。于是基因型的环境敏感性是它自己的值对环境值的回归。下面的例 8.3 将使这种方法更清楚。由基因型和特殊环境的互作引起的方差可从方差分析中估计出来，归因于敏感性差异的大小可从回归斜率的异质性中获得。其细节参见 Perkins 和 Jinks（1968）；另一个例子见 Zuberi 和 Gale（1976）。

例 8.3

（数据来自 J.L.Jinks 教授的热情提供）黄花烟草（*Nicotiana rustica*）的 10 个近交系分别生长于由不同种植密度和不同播种日期构成的 8 种特殊环境中。测量了每种环境中每个系的 8 棵植株的最终高度。方差分析表明基因型（系）之间和环境之间的差异都是显著的，并存在显著的基因型×环境互作。4 种基因型的环境敏感性绘制在图 8-2 中。为了估计一种基因型的环境敏感性，首先要将每一种环境的一般效应评估为在该环境中 10 种基因型的平均值。然后，每一种基因型值以环境平均值为背景绘图。回归直线的斜率度量了基因型的环境敏感性。回归系数 1.0 代表所有基因型的平均敏感性。图中 4 种基因型的敏感性显示在右端边缘处。中间的两种基因型具有特定的意义，它们在 10 个系中具有最高和最低的敏感性。它们在所有环境中具有几乎相等的平均值，但由于它们不同敏感性的结果，一个在好环境中较高，另一个在差环境中较高，优势的次序颠倒了（进一步的细节见 Mather 和 Jinks，1982，p.118）。

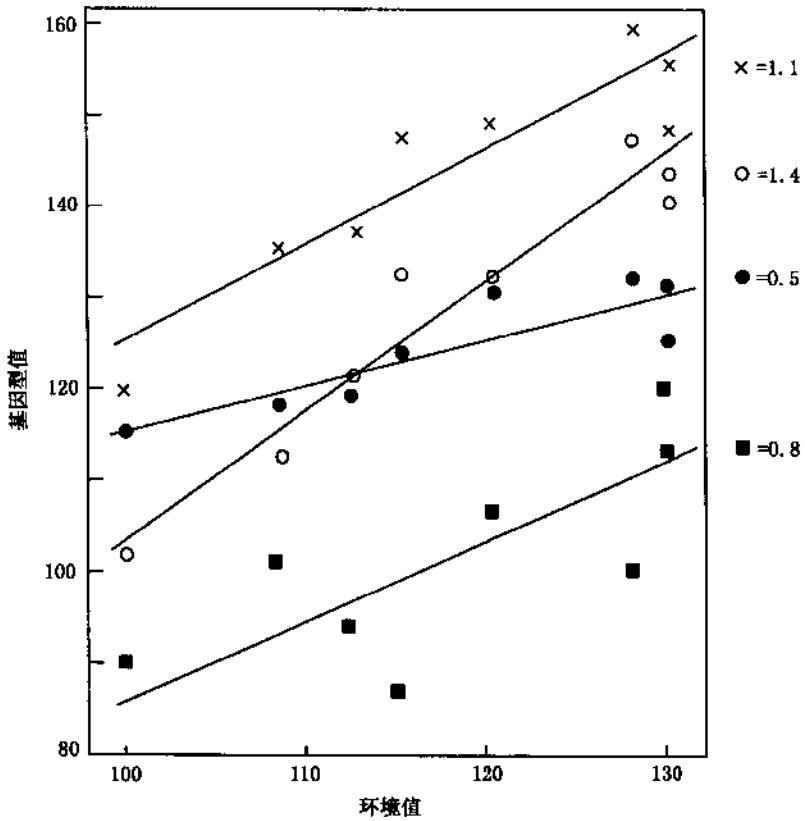


图 8-2 生长在 8 种特殊环境中的黄花烟草植株高度（cm），正如例 8.3 中解释的那样
（这些是在 Mather 和 Jinks，1982 的表 42 和 44 中编号为 3、6、7、10 的基因型。）

环境方差

根据定义，环境方差包含非遗传来源的所有变异，环境方差可由很不同的原因引起，

它的特性在很大程度上取决于所研究的性状和生物体。一般地说，环境方差是误差的一种来源，它降低了遗传研究的精确度，因此实验者或育种者的目的是通过细心管理或适当的实验设计来尽可能多地降低它。营养和气候因子是环境变异最普遍的外部原因，它们至少部分地处在实验控制之下。母体效应是形成环境变异的另一种来源，有时它是重要的，尤其是在哺乳动物中，但它对控制不那么敏感。母体效应是母亲对仔畜出生前和出生后的影响，主要是营养方面的：在下一章中涉及到亲属间的相似性时，我们将更多地谈到母体效应。度量的误差是变异的另一种来源，尽管它经常不很重要。当一个性状可用长度或重量单位测量时，它经常被测量得如此准确，以致于归因于测量的方差与方差的其余部分相比较是可以忽略的。可是，严格地讲，一些性状是不能被测量，而是必须根据判断分成等级。家畜的胴体品质就是一个例子。对这类性状，由测量引起的方差可能是相当大的。

除了诸如提到的那些可识别原因引起的变异之外，经常还有原因未知的巨大数量的非遗传变异，它们是不能通过实验设计来消除的。一般将之称为难以确定的变异。一些难以确定的变异可能由具有普遍意义上的“环境”条件引起，也就是说，由个体的外部条件引起——即使它们的属性是未知的。可是，一些难以确定的变异可能由发育变异所产生：也就是说，这种变异不能归因于外界的环境，而是在忽略其精确属性的情况下，归因于发育过程中的“偶然事件”或“误差”，以之作为一种一般性的原因。难以确定的变异主要是发育型变异的性状，是那些与解剖结构有联系的性状，在发育完成之后，它们不再变化，诸如果蝇（*Drosophila*）中的骨骼形状、色素沉着或刚毛数等。与之相反，对外部环境的影响较敏感的性状是那些与新陈代谢过程有关的性状，例如生长、繁殖和泌乳等。

例 8.4

人的出生重为研究有大量环境变异的性状提供了一个实例，它的变异的属性已被详细分析过（Penrose，1954；Robson，1955）。表 8-5 给出的表型方差分剖表明了所有可识别的变异来源的相对重要性，出生重被认为是孩子的一个性状。在出生重与出生前环境有关的意义上，所有的环境变异都是“母体的”，但母体环境的几种不同组分是可区别的。“母体基因型”，它解释了总表型方差的 20%，反映了母亲间在其子女出生重上的遗传变异（主要为加性）；即认为出生重是母亲的一个性状。“一般的母体环境”，它又解释了总表型方差的 18%，以同样方式反映了母亲间的非遗传变异。这两个组分总计 38%，是出生重变异的母体原因，对同一母亲所有子女的影响都是一样的。“个别的母体环境”指归因于母亲的原因，但在连续怀孕中的不同。同一性质的另两种原因——“母亲年龄”和“胎次”（即孩子是第一胎、还是第二胎，等等）——是可以分别识别的。最后，所有剩余的是难以确定的变异，它的原因不能识别。解释各种组分是如何估计出来的要占用太多的空间，直到第 10 章结束时才可恰当地去做。这里只须说一句，所有估计

表 8-5 新生儿体重的表型变异分剖

变异原因	总方差的%
遗传	
加性	15
非加性（近似）	1
性别	2
总的基因型	18
环境	
母体基因型	20
一般的母体环境	18
个别的母体环境	6
母亲年龄	1
胎次	7
难以确定的	30
总的环境	82

值都来自同卵双生、异卵双生、全同胞、姐妹的孩子以及其他种类表亲间相似程度的比较。应当注意到估计值不很精确，不可以拿来作为人类群体的确定参数。在不同数据的另一项研究中（Morton，1955），没有发现胎儿中的遗传因素影响，出生重的所有变异都是环境引起的。

人出生重的方差分剖。用总表型方差的百分率表示组分，见表 8-5。

多次度量：重复力

当对每一个体的性状作出一次以上的度量时，表型方差可分剖成个体内方差和个体间方差。这种分剖导致了称为重复力的方差组分比率，它有 3 个主要用途：显示重复度量的获益有多少，为比率 V_G/V_P 或 V_A/V_P 规定上限，据过去的记录预测将来的表现。它还可能阐明环境方差的属性。对应于重复力的方差分剖不是遗传理论的一个部分，因为分剖的是环境方差，而不是遗传方差。然而我们将看到，做这种分剖对遗传分析和育种方案确实有一些实际的影响。

有两种方法可以对一个性状的重复提供多次度量：通过时间上的重复和通过空间上的重复。产奶量和窝产仔数是在时间上重复的性状实例。产奶量可以在连续的泌乳期中度量，窝产仔数可以在连续的怀孕中度量。这样可以获得每一个体的若干次度量。于是，每个泌乳期产量的方差或每窝仔畜数目的方差可被分析成个体内的一种组分，它度量同一个体表现间的差异，以及个体间的一种组分，它度量个体之间的永久差异。个体内组分在来源上完全是环境的，是由连续表现之间环境的时间差异造成的。个体间组分部分是环境的，部分是遗传的，环境部分是由永久影响个体的环境条件引起的。因此，通过这种分析，可以把由时间环境条件引起的方差从剩余方差中分离出来，并且可以被度量。

在空间上重复的性状主要是结构上的或解剖上的性状，在植物中比在动物中更经常找到。例如，结一个以上果实的植物对果实的任一性状（例如它的形状或种子含量等）提供多于一个的度量值。动物中空间重复主要出现在身体两侧或一系列重复部位上度量的性状中，例如果蝇（*Drosophila*）腹节上的刚毛数。对空间重复的性状，其个体内方差在来源上仍然全部是环境的，但不同于时间上重复的性状的个体内方差，它代表由发育期间起作用的局部条件引起的“发育”变异。

为了可以一起讨论时间重复和空间重复，我们将使用术语*特殊环境方差* V_{Es}来指由暂时或局部条件引起的个体内方差；并使用术语*一般环境方差* V_{Eg}来指由个体间组分贡献的和由永久或非局部条件引起的环境方差。个体间组分占总表型方差的比率就是组内相关 r。它是同一个体重复度量之间的相关，被称为性状的*重复力*。这样通过重复力表示的表型方差可分剖成两个组分，V_{Es}对（$V_G + V_{Eg}$），重复力是：

$$r = \frac{V_G + V_{Eg}}{V_P} \qquad [8.12]$$

因此，重复力表示出单个度量值的方差比率，它归因于永久的或非局部的，既有遗传的又有环境的个体间差异。它允许单独估计由特殊环境引起的组分 V_{Es}，占总方差的比例，由下式给出：

$$\frac{V_{Es}}{V_P} = 1 - r \qquad [8.13]$$

从等式 [8.12] 可以看出，重复力为遗传决定程度 V_G/V_P 和遗传力 V_A/V_P 设定了上限。重复力比这两种比率的任何一个通常要容易确定得多，当不知道这两种比率时，重复力却能经常是已知的。遗传力是有实际重要性的一种比率，当然要比重复力小得多，而且它不可能较大，这种了解比一点不了解遗传力要好。随着性状的性质不同，当然也随着群体的遗传特性以及个体被饲养的环境条件不同，重复力也大不相同。表 8-6 中的估计值为在各种不同性状中可能发现的这类数值提供了一些概念，有两种情形在例 8.5 中进行了更详细的描述。

表 8-6 重复力的一些例子

	重复力
黑腹果蝇（*Drosophila melanogaster*）：	
腹部刚毛数（见例 8.5）	0.42
卵巢大小（见表 8-8）	0.54
小鼠：（原始数据）	
6 周龄体重	0.96
第 1 和第 2 窝的窝产仔数	0.45
绵羊：（Morley，1951）	
不同年份羊毛重量	0.74
牛（英国弗里生牛）：（Barker 和 Robertson，1966）	
第 1 和第 2 泌乳期的产奶量	0.40
第 1 和第 2 泌乳期的乳脂率	0.67

有两个假设暗含在重复力的概念中。第一个假设是不同度量的方差是相等的，并且以同一比例拥有它们的组分。第二个假设是不同的度量在遗传上反映了同一个性状——这一点将在第 19 章解释。如果这些假设不是正确的，那么重复力将变成一个有点模糊的概念，关于方差组分没有精确的意义。严格地说，表 8-6 中一些性状不符合这些假设。

例 8.5

腹节的腹部表面上的刚毛数是黑腹果蝇（*Drosophila melanogaster*）中被研究得很多的一个性状，因为技术上方便，而且它的遗传特性相对简单。我们已经几次提到过它，但还没有使用它作为一个例子。在雄性 3 个腹节的每一个上以及在雌性 4 个腹节的每一个上都大约有 20 根刚毛。因此每一腹节的刚毛数可作为一个空间上重复的性状来看待。Reeve 和 Robertson（1954）详细研究过这个性状的变异来源，所发现的方差组分列于表 8-7。

表 8-7 黑腹果蝇刚毛数的方差组分

		♂♂	♀♀
总表型方差	V_P	4.24	5.44
果蝇间方差	$V_G + V_{Eg}$	1.82	2.19
果蝇内方差	V_{Es}	2.42	3.25
重复力		0.429	0.403

对一个性状的重复力估计分离了由特殊环境引起的方差组分 V_{Es}，但它把环境方差的另一个组分——由一般环境引起的组分 V_{Eg}——留下与基因型方差混杂在一起，正如上面

例子显示的那样。为了完成分剖，我们需要将 V_{Eg} 从 V_G 中分出来。通过两种方式可以做到这一点：要么通过例 8.1 的方式估计基因型方差 V_G，要么通过一个遗传一致群体（例如近交系或 F_1 杂种）计算出重复力。在没有遗传变异的情况下，个体间方差组分仅由一般环境组分 V_{Eg} 组成，因而重复力度量出比率 $V_{Eg}/(V_{Eg}+V_{Es})$。对果蝇（*Drosophila*）的两个性状用这种方式分剖了环境方差，由此得到的全面分剖结果列在表 8-8 中。主要的兴趣点是，一般环境只产生了很小数量的变异。因此这些性状几乎不受外部环境的影响；或者，也许可以更准确地说，在消除不想要的环境变异来源方面饲养果蝇的技术是非常成功的。在实验条件下，实际上所有的非遗传变异都是由严格的局部原因引起的，这些原因独立地影响腹节或卵巢。因为 V_{Eg} 组分这样小，以致于重复力为遗传决定程度 V_G/V_P 提供了一个良好的估计值。而且，刚毛数的非加性遗传方差很小，因此重复力与遗传力 V_A/V_P 并没有很大差异。

表 8-8 黑腹果蝇两个性状的表型方差分剖（每一个组分以占单个度量总方差的百分率形式给出）

组分		(1) 刚毛数	(2) 卵巢大小
加性遗传	V_A	33	23
非加性遗传	V_{NA}	6	27
一般环境	V_{Eg}	3	4
特殊环境	V_{Es}	58	46
总表型	V_P	100	100

性状和数据来源：

(1) 位于一个腹节上的数目（Reeve 和 Robertson，1954）。分别计算雄性和雌性的结果，然后平均。

(2) 一个卵巢中卵巢管的数目（Robertson，1957*a*）。

这里的遗传组分的比例低于表 8-4 中的，因为表 8-4 指出的是两次度量之和的方差；见例 8.6。

多次度量的增益 了解重复力的一个有用方面是证明从重复度量中预测准确度的增益。重复力高，增益就小，反之亦然。问题是：与重复力相关联的准确度增益是什么？通过重复度量可被减少的唯一方差组分是由特殊环境引起的组分 V_{Es}，它减少的数量取决于进行度量的次数。假定每一个体被度量 n 次，这 n 次度量的平均值作为个体的表型值，记为 $P_{(n)}$。于是表型方差由基因型方差、一般环境方差和 $1/n$ 的特殊环境方差组成：

$$V_{P(n)} = V_G + V_{Eg} + \frac{1}{n}V_{Es} \qquad [8.14]$$

因此，增加度量次数可减少在表型方差中出现的由特殊环境引起的方差大小，而表型方差的这种下降代表了准确度的增益。n 次度量平均值的方差作为一次度量的方差比例可以用重复力的形式表示如下。根据等式［8.12］和［8.13］，用 r 和 $1-r$ 的形式书写组分，代入到等式［8.14］，给出：

$$V_{P(n)} = \left(r + \frac{1-r}{n}\right)V_P$$

重新整理导致：

$$\frac{V_{P(n)}}{V_P} = \frac{1+r(n-1)}{n} \qquad [8.15]$$

这种比率绘制在图 8-3 中，用来表明对有不同重复力的性状，其表型方差是如何因多次度量而降低的情况。当重复力高时，只有极少的特殊环境方差，多次度量对准确度给出极少的增益。当重复力低时，多次度量可导致准确度获得有价值的增益。可是，随着度量次数的增加，准确度的提高速度减少，很少值得进行两次或三次以上的度量。实践中，人们对度量的平均值或总和进行操作不会产生任何差异：虽然实际方差是不同的，但组分的相对大小不受影响。

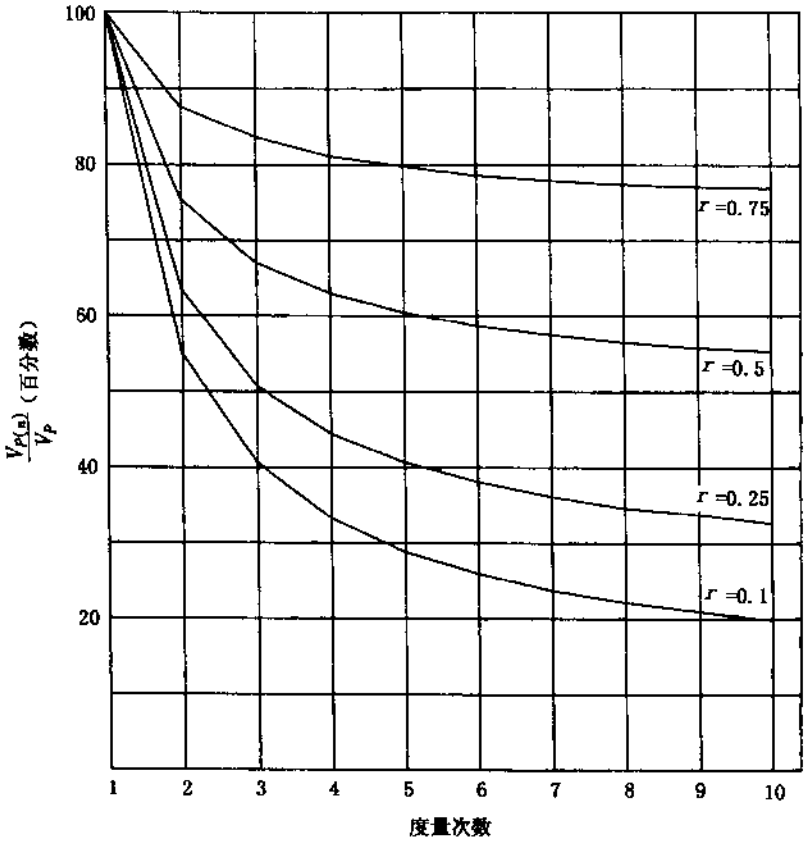


图 8-3 来自每一个体多次度量的准确度增益

垂直轴给出了 n 次度量平均值的方差与一次度量的方差的百分率。水平轴给出了最多达到 10 次的度量次数。4 条曲线指的是具有不同重复力的性状。

例 8.6

有关果蝇（*Drosophila*）腹部刚毛数的大部分研究是以两个节上的刚毛计数为基础的。表8-9显示了，通过计数两节而不是一节，方差组分的百分率是如何受影响的。栏目（1）给出了仅一节被计数时，表型方差组分的百分率，与表 8-8 中的一样。如果两节被计数，并且两次计数的平均值作为个体的表型，那么 V_E 将减少一半，但其他组分不变，这产生了栏目（2）中的数字。现在总方差 V_P 降低到 71。将每一种组分除以 71 再乘以 100 给出栏目（3）中的数字，即当表型值以两节计数为基础时方差组分的百分率。具有实际

重要性的一点是加性遗传方差从33%增加到46%（为什么 V_A 是46%而不是表8-4中的52%的理由是这两个估计值从不同品系中获得）。

表8-9 果蝇腹部刚毛数方差组分（以百分数表示）

	一个节	两个节	
	(1)	(2)	(3)
V_A	33	33	46.4
V_{NA}	6	6	8.2
V_{E_R}	3	3	4.1
V_{E_S}	58	29	41.3
V_P	100	71	100.0

育种方案从准确度增益中获得的益处是加性遗传方差的比例增加了。也就是说，两次以上度量的平均值比一单个度量值具有较高的遗传力，因而对个体育种值是一个更好的指导。可是，遗传力的这种增加不能依靠，除非早期提及的两个假设是正确的，即不同的度量有相等的方差并在遗传上代表同一性状。果蝇腹节上的刚毛数满足这些条件，因而在例8.6中得出的结论是正确的。假设不成立的一个性状是在连续泌乳期中母牛的产奶量（Rendel等，1957）。在这种情形中，几个泌乳期平均值的加性遗传方差的比例实际上低于仅用第一泌乳期的。

未来表现的预测 未来表现的预测是一个在很多章节中出现的问题。它没有一般的含义，但依赖于将方差分剖成归因于永久和暂时效应的组分，即通过重复力进行的分剖。过去和未来的表现必须以偏离过去和未来的群体平均值的离差形式来表示。良好的过去表现部分是个体受暂时环境影响而引起的，这些影响不能继续存在到随后的表现中，因此未来表现倾向于向群体平均值回归。如果不了解群体关于这两种表现的特性，例如第一和第二泌乳期的产奶量，那么是不进能行预测的。重复力，它是两种表现之间的相关，可告诉我们从对第一种表现的了解中能够怎样准确地预测第二种表现。预测本身是从第二种对第一种表现的回归系数中进行的。如果 x 和 y 分别是第一和第二种表现，$\bar{x}$ 和 $\bar{y}$ 是群体平均值，b 是 y 对 x 的回归系数，那么预测可由 $(y-\bar{y})=b\,(x-\bar{x})$ 给出。回归和相关系数之间的关系是 $b=r\,\sigma_y/\sigma_x$，其中 σ_x 和 σ_y 是标准差。

例8.7

可从Barker和Robertson（1966）关于英国弗里生母牛产奶量的数据说明未来表现的预测。这个数据涉及到具记录过第一、第二和第三个泌乳期产奶量的3 764头母牛。产奶量的平均值和标准差给在表8-10中，产奶量被转换成千克。在连续泌乳期中平均值和标准差两者都增加。让我们预测第一泌乳期产奶量为5 000kg的青年母牛的第二和第三个泌乳期的平均产奶量。需要的重复力是第二和第一泌乳期以及第三和第一泌乳期的相关；两者都是0.40。然后，可计算第二对第一泌乳期的回归，为 $b=r\,\sigma_2/\sigma_1$。第二个泌乳期的预测产奶量可从 $y-\bar{y}=b\,(x-\bar{x})$ 中获得，其中 y 是预测值，$\bar{y}$ 是第二泌乳期的平均产奶

量，x 是观察的产奶量为5 000kg，$\bar{x}$ 是第一泌乳期的平均产奶量。第二和第三泌乳期的计算和预测的平均值展示在表 8-10 中。

表 8-10 第二、第三泌乳期泌奶量的预测平均值

	泌乳期		
	1	2	3
平均值，kg	4 096	4 232	4 731
标准差（σ）	696	934	960
与第一泌乳期的相关（r）	—	0.40	0.40
对第一泌乳期的回归（b）	—	0.536	0.552
第一泌乳期的观察产奶量 = 5 000			
与平均值的离差 = +904			
第二泌乳期的预测产奶量 = 4 232 + (0.536 × 904) = 4 716.5			
第三泌乳期的预测产奶量 = 4 731 + (0.552 × 904) = 5 230.0			
第二和第三泌乳期中预测的平均值 4 973			

方差分剖的总结

本章已显示了一个遗传变异群体的表型方差如何被分剖成 4 个组分，其中两个是遗传的和两个是环境的。做这个需要三种数据，每一种再分剖成两个部分，但采用的方法不同。表 8-11 汇总了可做的不同分剖。

表 8-11 方差分剖的总结

所需数据	所做分剖	估计的比率
亲属间相似性	$(V_A):(V_{NA}+V_{Eg}+V_{Es})$	遗传力，V_A/V_P
遗传一致的群体	$(V_A+V_{NA}):(V_{Eg}+V_{Es})=(V_G):(V_E)$	遗传决定程度，V_G/V_P
多次度量	$(V_G+V_{Eg}):V_{Es}$	重复力 $(V_G+V_{Eg})/V_P$
所有 3 种	$V_A:V_{NA}:V_{Eg}:V_{Es}$	

习 题

8.1 在习题 6.1 中计算过烟草品种杂交 F_1 和 F_2 世代的叶片数目的方差。F_1 代的方差是 1.46，F_2 代方差为 5.97。估计 F_2 世代中的遗传决定程度。做这个必须作出什么假设？

[解答序号：7]

8.2 计算在（1）习题 7.2 和 7.5（所有 3 个群体），（2）习题 7.3 和 7.7，（3）习题 7.4 和 7.8 中指定群体涉及的基因所引起的加性遗传方差和显性方差的大小。

[解答序号：17]

8.3 当方差由具有下列显性度的一单个位点引起时，计算出总遗传方差中由显性引起的比例，即 V_D/V_G。（1）$d=\frac{1}{2}a$，即不完全显性，（2）$d=a$，即完全显性，（3）$d=2a$，即超显性。比率 V_D/V_G 取决于基因频率。绘图显示每一种情形的这个关系，并从图

中找出比率的最大近似值，以及这种情况发生时的基因频率是多大。

[解答序号：27]

8.4 涉及习题 7.9 和它的解答。计算出在指定群体中由两个基因引起的遗传方差的 3 个组分，即 V_A，V_D 和 V_I。将每一个组分以占总遗传方差 V_G 的百分率表示。

[解答序号：37]

8.5 来自一随机繁殖品系的样本为 10 的雌性小鼠，其第 1 和第 2 窝有下列活产仔数。

小鼠	1	2	3	4	5	6	7	8	9	10
第 1 窝	11	9	13	10	9	8	10	11	10	13
第 2 窝	10	12	12	10	8	6	12	9	12	12

（1）从这个样本中计算窝产仔数的重复力。

（2）从第 1 窝产仔数为（a）14 和（b）5 的同一品系中，你预测其他小鼠第 2 窝产仔数的期望大小将是多大？

[解答序号：47]

8.6 在猪生殖力的一项研究中，156 头母猪每头都有 10 窝仔猪，其窝产仔数的方差分析结果如下（“窝序”这一项指的是第一、第二、第三窝等窝仔猪数平均值之间的差异）。

来源	自由度	均方
母猪间	155	25.56
窝序间	9	93.95
母猪内	1.395	3.23

（1）计算窝产仔数的重复力。

（2）假定你正在计划一育种方案，期望具有高遗传力（V_A/V_P）。单个母猪的生殖力可通过一窝来度量，或者为有较大的精确度，也可通过几窝的平均值来度量。如果生殖力用初胎后 2 窝、初胎后 3 窝和初胎后 4 窝的平均值来度量，那么遗传力将会提高多少？假定重复力与从 10 窝中估计出来的相同。

数据来自 Olbrycht，T.M.（1943）*J. Agric. Sci.*，**33**，28 ~ 84。

[解答序号：57]

第 9 章　亲属间的相似性

亲属间的相似性是数量性状所表现的基本遗传现象之一，并且，相似度是对群体进行相对简单的度量而不用特殊实验技术就能确定的性状的一种特性。相似度提供了估计加性遗传方差大小的方法，并且正是加性方差的比例大小（即遗传力），主要决定了用于品种改良的最佳育种方法。因此，对亲属间相似性原因的了解是数量性状的实际研究及其在动植物改良中应用的基础。在这一章中，我们将探究亲属间相似性的原因，并从原理上说明加性方差的大小是怎样从观察的相似度中得到估计的，同时将估计遗传力的更实际的内容留在下一章予以考虑。

在上一章中，我们看到表型方差是怎样被分剖成归属于不同原因的组分。这些组分我们称之为方差的有因组分，并同以前那样用符号 V 表示。亲属间相似度的度量取决于用不同方法将表型方差分剖为相应于个体形成家系的组分。这些组分可由表型值直接估计，由于这个原因我们称之为表型方差的观察组分，用符号 σ^2 表示之，以使区别明显。例如，考虑将个体分成全同胞家系的情况。通过方差分析我们可将总观察方差分剖成两个组分，组间的和组内的。组间组分是各组“真实”均值与群体均值的方差，而组内组分则是个体与各组真实均值的方差。如果从很大量的个体中无误差地估计，一个组的真实均值是能够找到的。将在下一章中用例子给出这两个组分的估计解释。现在，亲缘个体间的相似性，即所讨论的全同胞间的相似性，可以被看作同一组中个体间的相似，或者看作不同组中个体间的差异。组内的相似性越大，组间的差异性也就越大。因此相似度可作为组间组分占总方差的比例来表示。这是组内相关系数，由下式给出：

$$t = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2}$$

其中，σ_B^2 是组间组分，σ_W^2 是组内组分（为避免与用符号 r 表示的其他相关发生混淆，习惯上用符号 t 表示表型值的组内相关）。组间组分表示同一组内成员共同的变异大小，并且它可同样地作为各组成员的协方差。在亲子间相似的情况下，观察值分组通常用对而不是组；一个亲本，或两个亲本的平均值，与一个后代或几个后代的平均值配对。对间方差组分只有在亲代值和子代值具有相同的方差时才有意义，但通常做不到这一点。因此，用亲子间的交叉乘积之和来计算亲子间的协方差，从而以子代对亲代的回归表示相似度。这里不用相关的原因在以后会明白。回归由下式给出：

$$b_{OP} = \frac{\mathrm{cov}_{OP}}{\sigma_P^2}$$

其中，cov_{OP} 是亲子间的协方差，σ_P^2 是亲本的方差。

这样，亲缘个体的协方差是群体的新属性，我们在探索不管是同胞还是子女与亲本等

亲属间相似性原因的过程中必须推导它。虽然协方差仅仅是总表型方差的一部分，并由上一章描述过的有因组分而组成，但在数量上和比例上视亲缘关系的类型而有所不同。通过查明有因组分怎样贡献协方差的情况，我们将看到怎样用一个观察的协方差来估计组成它的有因组分。

最普遍和最有用的亲缘关系是亲与子、半同胞、全同胞以及孪生子（在人类研究中）。这些亲缘关系中的协方差将得到充分的解释，其他亲缘关系的协方差在其后概述。孪生子有其特殊的问题，这将在下一章中涉及。

方差的遗传和环境来源都对亲属协方差有贡献，因而表型值的协方差等于遗传和环境协方差的和。遗传协方差将首先用它们引起的回归或相关来描述，而相似性的环境原因将在以后评述。

遗传协方差

现在我们的目标是要从理论上推导由遗传原因引起的亲属间的协方差，而暂时忽略可能使相似性存在的任何非遗传原因。这意味着我们必须推导亲缘个体基因型值的协方差。假设群体处于哈迪—温伯格平衡，所有亲本对所研究性状为随机交配。选型交配的效应将在下一章中考察。由位点间上位互作引起的任何方差开始时将被忽略，它的效应在以后简述。对于每一种亲缘关系，我们将描述两种推导协方差的方法，第一种较简明，第二种较详细。

后代与一个亲本

所要推导的协方差是个体的基因型值与其群体中随机交配产生的后代平均基因型值间的协方差。如果数值是用与群体均值的离差来表示的，则根据定义，后代平均值是亲代育种值的一半，正如第 7 章所解释的那样。因此要计算的协方差即为一个个体的基因型值与其育种值的一半的协方差，即 G 与 $\frac{1}{2}A$ 间的协方差。由于 $G = A + D$（D 为显性离差），则协方差是（$A + D$）与 $\frac{1}{2}A$ 间的协方差。将这些项乘在一起再对所有亲本求和，得：

$$\text{交叉乘积之和} = \sum \frac{1}{2}A(A + D)$$

$$= \frac{1}{2}\sum A^2 + \frac{1}{2}\sum AD$$

两边都除以亲本个数即得到协方差 $\frac{1}{2}V_A + \frac{1}{2}\text{cov}_{AD}$。在前一章中已证明 cov_{AD} 为 0，那么：

$$\text{cov}_{OP} = \frac{1}{2}V_A \qquad [9.1]$$

因此，后代与一个亲本的遗传协方差是亲代加性遗传方差的一半。

第二种得到协方差的较详细的方法是考察单个位点的效应。这将通过含有两个等位基因的一个位点来做，但结论对具有任何数目等位基因的位点同样有效。表 9-1 给出了亲代的基因型、它们在群体中的频率以及用它们与群体均值离差表示的基因型值（来自表 7-9）。右边栏给出了后代的平均基因型值，它们是表 7-9 给出的亲代育种值的一半。于是后代与一个亲本的协方差等于平均交叉乘积，这可通过将三栏乘在一起——频率 × 亲代基因

型值×子代基因型值——再对亲代3种基因型求和而获得。将含 α^2 和 αd 的项合并后，我们得到：

表9-1 亲代基因型及其频率、基因型值

亲代			后代
基因型	频率	基因型值	平均基因型值
A_1A_1	p^2	$2q\ (\alpha - qd)$	$q\alpha$
A_1A_2	$2pq$	$(q-p)\ \alpha + 2pqd$	$\frac{1}{2}\ (q-p)\ \alpha$
A_2A_2	q^2	$-2p\ (\alpha + pd))$	$-p\alpha$

$$\begin{aligned}\mathrm{cov}_{OP} &= pq\ \alpha^2\ (p^2 + 2pq + q^2)\ + 2p^2q^2\alpha d\ (-q + q - p + p)\\ &= pq\ \alpha^2\\ &= \frac{1}{2}V_A\end{aligned}$$

这是因为根据等式 [8.3a]，$V_A = 2pq\ \alpha^2$ 之故。对所有位点求和，我们又得到后代与一个亲本的协方差等于加性方差一半的结论。

子代对一个亲本的回归是通过将协方差除以亲本的方差即群体表型方差而得到的。这样，回归为：

$$b_{OP} = \frac{1}{2}\frac{V_A}{V_P} \qquad [9.2]$$

以上通过考察每一亲本后代的平均值推导出协方差，但没有指明平均值基于的后代数目。实际上，不管后代数目是多少，即使只使用一个，协方差总是相同的，其原因如下。n 个后代的均值是 $(1/n)\ \sum O$，其中 $\sum O$ 是后代值的总和。一个亲本与 n 个后代均值的协方差是 $\mathrm{cov}[P,(1/n)\sum O] = (1/n)\sum \mathrm{cov}(P,O) = \mathrm{cov}(P,O)$，这是亲本与任一后代的协方差。这个结论适用于任何种类的亲属关系。因此总的来说，任何个体和多个亲属平均值的协方差与其和这些亲属间任何一个的协方差是相等的。由于后代方差不参与回归的计算，因而子代对亲代的回归也不受所用后代数量的影响。

后代与中亲值

后代均值与双亲均值（通常称为“中亲值”）的协方差可以用以下方法推导。以O代表子代均值，P和P′代表双亲值。则中亲值 $\bar{P} = \frac{1}{2}(P + P')$。交叉乘积之和是 $\sum O\bar{P} = \frac{1}{2}(\sum OP + \sum OP')$，协方差为 $\mathrm{cov}_{O\bar{P}} = \frac{1}{2}(\mathrm{cov}_{OP} + \mathrm{cov}_{OP'})$。如果P和P′具有相同的方差，那么 $\mathrm{cov}_{OP} = \mathrm{cov}_{OP'}$，因此：

$$\mathrm{cov}_{O\bar{P}} = \mathrm{cov}_{OP} = \frac{1}{2}V_A \qquad [9.3]$$

这样，只要两性具有相同的方差，子代与中亲值的协方差就和子代与一个亲本的相同，我们已经知道后者等于加性方差的一半。

证明子代与中亲值的协方差的较详细方法是相当费力的，但必须将它推导出来，因为

获得全同胞的协方差将需要它。但我们将省略代数化简的一些步骤。表 9-2 是用与子代和一个亲本同样方式制成的，但现在我们必须将交配类型及它们的频率列成表，以代替单一亲本。与每种交配类型相对，我们填入双亲平均基因型值，即中亲值；然后是 3 种基因型在后代之间的比例以及后代的平均基因型值。以 a 和 d 的形式而不用与群体均值离差形式来书写基因型值将使这项工作做起来容易些。表的最后两栏给出了后代均值与中亲值的乘积以及后代值的平方，供以后使用。要得到后代均值与中亲值的协方差，我们取后代均值与中亲值的乘积，再把它乘以交配类型的频率，然后对所有交配类型求和。这得出平均乘积（MP），由于这里的数值不是以与均值的离差表示的，我们须从中减去群体均值的校正项。由于亲代与子代的均值是相等的，因而校正项简单地就是群体均值的平方（M^2）。MP 和 M^2 都含有 a^2、ad 和 d^2 的项。对这些项进行合并稍作简化，得到：

$$MP = a^2[p^3(p+q) + q^3(p+q)] + 2adpq(p^2 - q^2) + d^2pq(p^2 + 2pq + q^2)$$
$$M^2 = a^2(p^2 - 2pq + q^2) \qquad + 4adpq(p-q) \quad + 4d^2p^2q^2$$

然后，当对所有位点求和时，有：

$$\begin{aligned}\mathrm{cov}_{O\bar{P}} &= MP - M^2 \\ &= a^2pq - 2adpq(p-q) + d^2pq(p-q)^2 \\ &= pq[a + d(q-p)]^2 \\ &= pq\alpha^2 \\ &= \frac{1}{2}V_A\end{aligned}$$

虽然子代与双亲均值的协方差和其与一个亲本的协方差相同，但相似度却不同。子代对中亲值的回归为 $b = \mathrm{cov}_{O\bar{P}}/\sigma^2_{\bar{P}}$，其中 $\sigma^2_{\bar{P}}$是中亲值的方差。因为总的来说，n 个个体均值的方差是单个个体方差的 n 分之一，如果两性方差相等，那么 $\sigma^2_{\bar{P}} = \frac{1}{2}V_P$。因此子代对中亲值的回归为：

$$b_{O\bar{P}} = \frac{\frac{1}{2}V_A}{\frac{1}{2}V_P} = \frac{V_A}{V_P} \qquad [9.4]$$

它是对一个亲本回归的两倍。同一个亲本的一样，所用后代数目并不影响协方差或回归。

由于仅仅与方差的有因组分有关，子代对亲本的回归是相似度的一个有用度量。但子代与亲本间的相关就没有这种有用特性。因为相关用$\mathrm{cov}_{OP}/(\sigma_O\ \sigma_P)$计算，其中 σ_O 和 σ_P 分别为子代和亲代方差的平方根，而不管是用单个数值还是用多个个体的均值计算的。因此相关受后代数目及亲代数目多少的影响。如果只有一个后代，则其与一个亲本的相关就与回归相同，但在其他情形下，相关和回归是不同的。一个后代与中亲值的相关是 $\left(\sqrt{\frac{1}{2}}\right)V_A/V_P$。当有多于一个的后代时，相关就取决于后代观察均值的方差，并且与方差的有因组分没有简单的关系。

半同胞

半同胞是有一个共同的亲本而另一个亲本不同的个体。因此，一组半同胞是一个个体

表 9-2　交配类型、中亲值及后代平均基因型值

亲本基因型		交配频率	中亲值	后代			后代平均值	后代均值×中亲值	后代均值平方
				A_1A_1 a	A_1A_2 d	A_2A_2 -a			
A_1A_1	A_1A_1	p^4	a	1	—	—	a	a^2	a^2
A_1A_1	A_1A_2	$4p^3q$	$\frac{1}{2}(a+d)$	$\frac{1}{2}$	$\frac{1}{2}$	—	$\frac{1}{2}(a+d)$	$\frac{1}{4}(a^2+2ad+d^2)$	$\frac{1}{4}(a^2+2ad+d^2)$
A_1A_1	A_2A_2	$2p^2q^2$	0	—	1	—	d	0	d^2
A_1A_2	A_1A_2	$4p^2q^2$	d	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}d$	$\frac{1}{2}d^2$	$\frac{1}{4}d^2$
A_1A_2	A_2A_2	$4pq^3$	$\frac{1}{2}(-a+d)$	—	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}(-a+d)$	$\frac{1}{4}(a^2-2ad+d^2)$	$\frac{1}{4}(a^2-2ad+d^2)$
A_2A_2	A_2A_2	q^4	$-a$	—	—	1	$-a$	a^2	a^2

与一个异性随机群体交配，并且每次交配只产生一个子女的后代。因此根据定义，半同胞组的平均基因型值就是共同亲本育种值的一半。协方差是半同胞组的真实均值方差，因此是共同亲本一半育种值的方差，它是加性方差的 1/4。

$$\mathrm{cov}_{(HS)} = V_{\frac{1}{2}A} = \frac{1}{4}V_A \qquad [9.5]$$

这个协方差也可由表 9-1 已给出的值用较详细的方法证明。协方差是右边栏里列出的后代组均值的方差。将后代值平方，并乘以其频率得到：

半同胞家系平均值的方差

$$= p^2q^2\alpha^2 + 2pq\cdot\frac{1}{4}(q-p)^2\alpha^2 + q^2p^2\alpha^2$$

$$= pq\alpha^2\left[pq + \frac{1}{2}(q-p)^2 + pq\right]$$

$$= pq\alpha^2\left[\frac{1}{2}(p+q)^2\right]$$

$$= \frac{1}{2}pq\alpha^2$$

由于 $2pq\alpha^2 = V_A$（由等式[8.3a]），因此对所有位点求和有：

$$\mathrm{cov}_{(HS)} = \frac{1}{4}V_A$$

同胞间相似度用组内相关表示，而组内相关是组间方差即协方差与总方差的一个比例。因此半同胞的相关是：

$$t = \frac{1}{4}\frac{V_A}{V_P} \qquad [9.6]$$

全同胞

因为显性方差的贡献，所以全同胞的协方差比目前已考察的亲缘关系的协方差要复杂一些。首先考察仅由加性方差造成的协方差。全同胞有两个共同的亲本，于是全同胞组的平均基因型值等于双亲的平均育种值。用 A 和 A' 代表双亲的育种值。那么，协方差是 $\frac{1}{2}(A+A')$的方差，如果两性的加性方差相同，则$\frac{1}{4}(V_A + V_A') = \frac{1}{2}V_A$。现在考察显性的贡献。很容易想到协方差用随机抽取的同胞对的交叉乘积和来计算。设亲本有基因型 A_1A_2 和 A_3A_4。于是后代有 4 种基因型 A_1A_3、A_1A_4、A_2A_3 和 A_2A_4，每一种的频率为$\frac{1}{4}$。设所选第一个同胞有这些基因型的任何一种。那么第二个同胞有相同基因型的可能性为$\frac{1}{4}$。这样所有同胞对的$\frac{1}{4}$有相同的基因型，因此也有相同的显性离差 D。对有相同显性离差的这些对子，它们的显性离差的交叉乘积为 D^2；有不同显性离差的其他对子，其平均交叉乘积为 0。因此，对于所有对子，其交叉乘积之和为$\frac{1}{4}\sum D^2$，这样平均交叉乘积为$\frac{1}{4}V_D$。这是由显性离差造成的协方差，并且加进了由育种值造成的协方差中。因此全同胞的遗传协方差为：

$$\mathrm{cov}_{(FS)} = \frac{1}{2}V_A + \frac{1}{4}V_D \qquad [9.7]$$

获得全同胞协方差的第二种方式可由表 9-2 得出，但要有一点附加工作。协方差是家系平均值的方差。右边栏里给出了后代均值的平方，可以看出这些正好与后代平均值和中亲值的乘积相同，除了中间含有 d^2 项的两项不相等之外。因此均方（MS）可从已算出的平均乘积（MP）中得到，这样：

$$MS = MP + d^2 2p^2q^2 - \frac{1}{4}d^2 4p^2q^2$$

$$= MP + d^2p^2q^2$$

均值的校正与以前的相同，因此我们有：

$$\mathrm{cov}_{(FS)} = \mathrm{cov}_{\overline{OP}} + d^2p^2q^2$$

$$= pq\alpha^2 + d^2p^2q^2$$

由于 $2pq\alpha^2 = V_A$（来自等式[8.3a]）以及 $4d^2p^2q^2 = V_D$（来自等式[8.4]），因此对所有位点求和，全同胞的协方差为：

$$\mathrm{cov}_{(FS)} = \frac{1}{2}\mathrm{V_A} + \frac{1}{4}\mathrm{V_D}$$

全同胞的相关为：

$$t = \frac{\frac{1}{2}V_A + \frac{1}{4}V_D}{V_P} \qquad [9.8]$$

由于 $\mathrm{cov}_{(FS)} - 2\mathrm{cov}_{(HS)} = \frac{1}{4}\mathrm{V_D}$，因而在原理上，全同胞与半同胞协方差间的差异提供了一种估计显性方差的方法。但实际上，只有当表型协方差中无环境贡献时才可以这样做。

孪生子

孪生子的遗传协方差是很简单的。异卵双生子与全同胞相同，它们的遗传协方差就等于全同胞的协方差。同卵双生子具有相同的基因型，因此在对子内没有遗传方差，并且总遗传方差表现在对子间组分内。因此遗传协方差为：

$$\mathrm{cov}_{(MZ)} = \mathrm{V_G} \qquad [9.9]$$

概 要

从上面所解释的关系中可以看出，协方差是由方差有因组分 V_A 和 A_D 的简单分数（fraction）构成的，在所涉及的情形中该分数为 $\frac{1}{2}$ 或 $\frac{1}{4}$。这些分数或系数以一简单方式与亲属及其亲代的共祖率相关，因此任何类型亲属的协方差可从适当谱系的考察之中容易地推导出来。用 r 表示协方差中的加性遗传方差的分数，u 为显性方差的分数，那么任何类型亲缘关系的一般协方差为：

$$\mathrm{cov} = rV_A + uV_D \qquad [9.10]$$

用 P 和 Q 代表需要求得协方差的具有亲缘关系的两个个体，A、B 和 C、D 分别为其双亲，如图 5-4 所示。然后以 f 代表第 5 章中所说明的共祖率，可得 r 和 u 的值为：

$$r = 2f_{PQ} \qquad [9.11]$$

$$u = f_{AC}f_{BD} + f_{AD}f_{BC} \qquad [9.12]$$

(见 Crow 和 Kimura,1970,p.134)。表 9-3 给出了 r 和 u 的值,表中汇总了已描述过的亲缘关系并添加了一些其他更远的亲缘关系。等式[9.11]和[9.12]适用于随机交配群体。可是,亲本的近交并不影响它们,但如果亲属自身是近亲的,则有:

$$r = 2f_{PQ}/\sqrt{(1+F_P)(1+F_Q)}$$

(Crow 和 Kimura,1970,p.138)。

表 9-3 亲属协方差中方差组分的系数

亲缘关系		系数 V_A 的 r	系数 V_P 的 u
同卵双生		1	1
第一级	子代与亲代	$\frac{1}{2}$	0
	全同胞	$\frac{1}{2}$	$\frac{1}{4}$
第二级	半同胞	$\frac{1}{4}$	0
	后代与祖代	$\frac{1}{4}$	0
	叔(姨)与侄(姪)	$\frac{1}{4}$	0
	双重表亲	$\frac{1}{4}$	$\frac{1}{16}$
第三级	后代与曾祖代	$\frac{1}{8}$	0
	单重表亲	$\frac{1}{8}$	0

加性方差的系数 r 有时称为所讨论亲属间的*亲缘系数*,或*理论相关*。它是亲属育种值间的相关,它代表如果所有表型方差全是加性遗传时所要求得的相关。在下一章中考察遗传力估计时,我们将回到这一点上。显性方差的系数 u 代表亲属通过同一血统具有相同基因型的可能性。除非亲缘个体通过它们各自双亲具有像全同胞及双重表亲那样的共祖率通径,否则 u 值为 0。

上位互作

由位点间上位互作引起的方差只有小部分贡献给亲属协方差。在第 8 章中已经注意到,按照互作位点数以及互作是在育种值间还是在显性离差之间,将互作方差 V_I 再细分为各种组分。当包含互作组分时,任何类型亲属的一般协方差如下(细节见 Kempthorne,1955*a*,*b*):

$$\left.\begin{aligned}\text{cov} = {} & rV_A + uV_D + r^2V_{AA} + ruV_{AD} + u^2V_{DD} \\ & + r^3V_{AAA} + r^2uV_{AAD} + ru^2V_{ADD} + u^3V_{DDD} \\ & + \cdots\end{aligned}\right\} \qquad [9.13]$$

表 9-4 给出了含有两因子互作的协方差中方差组分的系数。子代—亲代协方差同样适用于一个亲本和中亲值。考察互作组分所得结论如下。首先,只有小部分互作组分贡献给任一协方差,最大值为 1/4;多于两位点之间的互作贡献更小。其次,除非显性方差本身有贡献,

否则包含显性方差的互作组分也没有贡献。第三，育种值的互作，V_{AA}、A_{AAA}等，对所有亲属协方差都有贡献。在第7章中，我们注意到，如果位点间存在互作，则那里给出的育种值的两种定义是不等价的。现在我们可以明白这是怎样产生的。一种定义（理论定义）是单独以每一个位点上基因的平均效应为依据的，它不包含互作离差。另一种定义（实用定义）是以后代的度量值为依据的，它包括育种值互作离差的部分。类似地，加性方差，即育种值的方差，也有两种定义。理论定义，前一章已描述过，是对单独考察的位点求和而获得的，它不包含互作方差。实用定义是它决定了亲属的协方差，并从中估计而来，因此它含有加性×加性互作方差的部分；例如，后代—亲本协方差的两倍估计了 $V_A + \frac{1}{2}V_{AA}$ + 较小部分的较高阶互作，其中 V_A 是理论定义的加性方差。估计加性方差的唯一方法来自亲属的协方差，例如对选择反应的预测正需要这一点。因此，当我们谈到"加性方差"或书写"V_A"时，将这些视为包括互作方差部分的实用定义是最好的。可是，实际上，这两种定义之间的差异与估计误差相比，可能是极其微不足道的。

从表9-4中的系数可以看出，在原理上互作组分是如何被估计的。例如，$\mathrm{cov}_{OP} - 2\mathrm{cov}_{(HS)} = \frac{1}{8}V_{AA}$。可是，要估计互作组分，需要高度精确的复杂实验，以提供排除了环境协方差的许多不同类型亲属间协方差的比较。这在动物中几乎是不实际的，但在植物中是可行的。实例可见 Pooni 等（1978）和 Chi 等（1969），他们发现玉米7个性状中存在可忽略的互作方差。

表9-4 包含两因子互作贡献的亲属协方差

亲 属		方差组分及其贡献系数				
		V_A	V_D	V_{AA}	V_{AD}	V_{DD}
后代—亲代	cov_{OP} =	$\frac{1}{2}$	—	$\frac{1}{4}$	—	—
半同胞	$\mathrm{cov}_{(HS)}$ =	$\frac{1}{4}$	—	$\frac{1}{16}$	—	—
全同胞	$\mathrm{cov}_{(FS)}$ =	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$
一般	cov =	r	u	r^2	ru	u^2

连锁效应 在获得以上给出的所有协方差的过程中，我们忽略了连锁效应，对所有位点求和有一个暗示的假设，即所有位点都独立分离。如果群体处于连锁平衡，则连锁不影响协方差，除非位点间存在上位互作；全同胞和半同胞的协方差则由于互作位点间的连锁而增加。具体细节见 Cockerham（1956）以及 Weir、Cockerham 和 Reynolds（1980）。

环境协方差

遗传原因不是亲属间相似性的唯一原因；也还有环境因素使得亲属趋于彼此相似：某些类型的亲属比其他的更相似。如果一个家系的成员在一起饲养，如同人类家庭或猪、小鼠的一窝仔等，它们生活在共同的环境里。这意味着引起非亲属个体间差异的一些环境因素，不是同一家系中成员间差异的原因。换句话说，有一个环境方差的组分，它对家系均值间的方差有贡献，但不贡献给家系内的方差，因此，它对亲缘个体的协方差有贡献。这个*组间环境*组分，我们用符号 V_{Ec} 表示，通常称之为*共同环境*，当我们把这个组分看作一个组中成员间

的相似原因时，与我们把它看作不同组成员间的差异原因时相比，这个术语显得更合适。环境方差的剩余部分，以 V_{Ew} 表示，它由差异性原因产生，与个体是否有亲属关系无关。因此它出现在方差的组内组分中，但不给组间组分做贡献，组间组分是各组真实均值的方差。因此，在考察亲属相似性的过程中，环境方差必须分剖成两个组分：

$$V_E = V_{Ec} + V_{Ew} \qquad [9.14]$$

V_{Ec}，作为环境方差的组分之一，对亲属个体的协方差有贡献。

共同环境方差的来源很多且各种各样，由诸如营养、气候条件或在人类还有文化影响等环境因素引起。无论何时家系在这些因素方面有差异，由环境引起的家系均值间差异都存在，并作为 V_{Ec} 组分出现在协方差中。当我们估计表型方差的观察组分时，我们所指的 V_{Ec} 组分取决于个体的分组方式。不论分析的方式是什么，组均值间的方差可归因于环境原因的部分就称为 V_{Ec} 组分。因此，这个组分的特性就取决于所用的分析方式。如果分析中所用的组为全同胞家系，那么 V_{Ec} 组分表示全同胞间相似性的环境原因；如果组为半同胞，它表示半同胞间相似性的原因。在亲子亲缘关系中，一个可比的协方差项表示亲子间相似的环境原因。因此，无论我们何时度量一个表型方差以用于估计方差的一个有因组分，我们必须确定它是否包含一个由共同环境引起的可测量组分，而这往往是根据对生物和性状的生物学理解作出判断的事。在实验中，通过适当设计往往可以消除大部分的 V_{Ec} 组分。例如，同一家系的成员不必总饲养在同一小瓶、笼子或地点：它们在各种饲养环境中可以随机化配置。或者，将家系分成两个或更多的组，这样 V_{Ec} 组分就可度量并从协方差中减去。

母体效应 母体效应是一个常见的并常引起麻烦的环境相似性的来源，尤其对哺乳动物。仔畜在它们生命的开始阶段受制于母体环境，这影响到甚至在成年时度量的许多数量性状的表型值，从而引起同一母亲的后代彼此相似。有两种母体效应需要分清。第一种，所研究性状的母亲表型值可以影响后代同一性状的数值。例如，体格大的鼠比体格小的鼠能提供较多的乳，因此它们的子代生长较好。这导致了一个由环境引起的子代体重与母亲体重间的相似性。更进一步，同一母亲的后代因为分享同一奶源而在体重上彼此相似。因此，这类母体效应贡献一个环境组分给子代与母亲的协方差及全同胞或母系半同胞的协方差。Falconer（1965*b*）描述了评价母亲表型对其后代同一性状影响的一种方法。第二种母体效应引起同一母亲后代间的相似，但不引起子代与其母亲间的相似。当引起母体效应的母亲性状不是正在研究其协方差的性状时，就引起了这种母体效应。例如，幼鼠尾巴的生长受窝中温度的影响。母亲们照料后代并不都是兢兢业业的，因而导致家系间窝内温度有差异。这在同胞协方差上产生了关于尾巴长度的一个环境组分。但是窝内温度与母亲尾巴长度无关，因此子代与母亲间在尾长上没有环境协方差。归因于母体效应的后代间变异起源于引起母体效应的那个性状在母亲间的变异，例如乳汁产量。母体性状在一定程度上由母亲基因型决定。因此在子代中见到的环境方差 V_{Ec} 在某种程度上是母亲中一些其他性状遗传变异的结果。将母体效应的遗传基础纳入考察时，亲属间的相似性会变得非常复杂。细节见 Willham（1963，1972）和 Thompson（1976）。

各种类型的亲属可能受制于相似性的一个环境来源。可是在以下接着要讲的内容中，我们将作出一些简化，除全同胞外，对所有其他亲属关系都忽略 V_{Ec} 组分。全同胞的共同母体环境通常是最难通过试验设计克服的环境相似来源。因而，一个 V_{Ec} 组分对全同胞协

方差的贡献比对任何其他类型亲属协方差的贡献机会更多且数量更大。

竞争 对于环境原因还必须简单提及一个途径，通过它亲属间的相似性可被减小而不是增加。这发生在同一家系成员争夺有限资源时。例如，假设动物同胞家系按每一家系饲养在一个单独畜圈内，所有圈内给以固定的相同数量的饲料，所考察的性状为生长速率。那么家系间在生长速率上将很少有变异或没有变异，并且协方差会因之减小。可是，家系内仍将有变异；的确，因为竞争是一个额外的变异来源，所以可能比不限制饲喂有更大的变异。因此组内相关会因对固定资源的竞争而减小。因为如果一个体得到较多食物，另一些个体必定得到的要少，所以相关甚至是负的。在植物中，竞争是一个重要的因子，它通常使同胞相关无意义，尤其是对与产量有关的性状。

表型相似性

表型值的协方差是由遗传和环境原因引起的协方差之和。这样将前面两个部分的结论放在一起，我们得到汇总在表9-5的表型协方差（要记住，环境协方差的一些可能来源被忽略了，尤其在包含母亲的亲子关系中）。在除全同胞以外的所有这些亲缘关系中，协方差为加性遗传方差的1/2或1/4。通过观察亲属的表型协方差，我们可以估计加性遗传方差的大小。与之相似，回归或相关提供了估计加性遗传方差数量比例 V_A/V_P，即遗传力的一种方法，这是度量亲属间相似度的主要用途。在下一章中，我们将更全面地考察估计遗传力的方法。

表9-5 亲属间的表型相似性

亲　属	协方差	回归（b）或相关（t）
子代与一个亲本	$\frac{1}{2}V_A$	$b=\frac{1}{2}\frac{V_A}{V_P}$
子代与中亲	$\frac{1}{2}V_A$	$b=\frac{V_A}{V_P}$
半同胞	$\frac{1}{4}V_A$	$t=\frac{1}{4}\frac{V_A}{V_P}$
全同胞	$\frac{1}{2}V_A+\frac{1}{4}V_D+V_{Ec}$	$t=\frac{\frac{1}{2}V_A+\frac{1}{4}V_D+V_{Ec}}{V_P}$

习　题

9.1 一对同卵双生子与无亲缘关系的配偶结婚所生的孩子间的亲缘系数 r 是多大？

［解答序号：8］

9.2 习题7.4、7.8和8.2（3）涉及一个特定群体中两个基因对小鼠被毛色素沉着的影响。假设个体的度量受到环境方差的作用，其数量达到总遗传方差的1/3，但家系均值间无环境差异。那么在此群体中，下列亲属关系间的表型相似性是多大：（1）子代与中亲，（2）子代与一个亲本，（3）全同胞，（4）半同胞，（5）双重表亲？

［解答序号：18］

9.3 下列是已报道过的总指脊数的相关。它们互相之间是一致的吗？在比较相关时必须作出什么假设？

中亲—子： $r = 0.69 \pm 0.03$

父—子： $r = 0.50 \pm 0.04$

母—子： $r = 0.49 \pm 0.04$

数据来自 Holt，S.B.（1956/57）*Acta Genet.*，**6**，473～476。

[解答序号：28]

9.4 下面的估计值指的是在不同实验程序下小鼠的窝产仔数。窝产仔数即出生的仔畜数，是母亲的一个性状。一种情况是，通过移去超过 8 个的仔畜或以其他窝抚养的仔畜补足 8 个，将母亲窝产仔数度量过的窝数都标准化为出生时 8 个仔畜。另一种情况是，对窝数不作处理并将出生时的大多数仔畜饲养至断奶。除抽样误差及品系差异外，提出可解释估计值间差异的原因。两种情况下，同一窝的生母是全同胞。

	标准化的窝仔畜	未标准化的窝仔畜
全同胞相关	0.055	0.107
女儿—母亲回归	0.045	-0.028

数据来自 Eisen，E.J.（1978）*Genetics*，**88**，781～811；Falconer，D.S.（1965）pp. 763～774 in Geerts，S.J.（ed.），*Genetics Today*. Proc. XI Internat.Congr.Genetics，Vol.**3**. Pergamon，Oxford。

[解答序号：38]

第 10 章　遗传力

数量性状的遗传力是它最重要的特性之一。正如我们所看到的，遗传力表示总方差中归因于育种值差异的比例，而这就是决定亲属间相似程度的部分。但是，在数量性状的遗传学研究中，遗传力最重要的功能是其预测作用，指示表型值作为育种值指南的可靠性。只有个体的表型值可直接度量，但正是育种值决定它们对下一代的影响。因此，如果育种者或实验者根据表型值选择个体作为亲本，那么他在改变群体特性方面的成功只能从对表型值和育种值之间一致性程度的了解中来预测。这种一致性程度可通过遗传力来度量，正如下面的考察所看到的那样。

遗传力定义为加性遗传方差与表型方差的比率：

$$h^2 = \frac{V_A}{V_P} \qquad [10.1]$$

(习惯上用符号 h^2 表示遗传力本身，不表示它的平方。这个符号源自 Wright(1921)的术语，其中 h 表示对应的标准差的比率。)遗传力的一个等价意义是育种值对表型值的回归：

$$h^2 = b_{AP} \qquad [10.2]$$

这种意义的等价性可从类似于在前一章中获得后代与单亲的遗传协方差的推导中看出来。如果我们将表型值分剖成育种值以及由环境、显性和互作离差组成的残余值(R)，那么 $P = A + R$。因为 A 和 R 不相关，$\mathrm{cov}_{AP} = V_A$，因此 $b_{AP} = V_A/V_P = h^2$。

我们可能还注意到，育种值和表型值之间的相关 r_{AP} 等于遗传力的平方根。这可从相关和回归系数之间的一般关系中得出，它给出：

$$r_{AP} = b_{AP}\frac{\sigma_P}{\sigma_A} = h^2\frac{1}{h} = h \qquad [10.3]$$

通过将遗传力看作是育种值对表型值的回归，我们可看出一个个体的估计育种值是其表型值和遗传力的乘积：

$$A_{(期望)} = h^2 P \qquad [10.4]$$

育种值和表型值两者都可计算为与群体平均值的离差。换句话说，遗传力表示了表型值作为育种值指南的可靠性，或者表示了表型值和育种值之间一致性的程度。由于这个原因，遗传力进入到几乎每一个与育种方法有关的公式中，并且很多关于方法的实际决策取决于它的大小。可是，这些问题将在下一章中考虑；这里我们仅关切地指出遗传力的确定是数量性状遗传学研究中的首要目标之一。

认识到遗传力不仅是一个性状的，而且是群体的，个体所遭遇的环境条件的，表型度量方式的一种特性是重要的。既然遗传力的值取决于所有方差组分的大小，那么这些组分中任何一个的变化都会影响它。所有遗传组分都受基因频率的影响，而且根据群体过去的历史，遗传组分可能在各个群体间有所不同。特别是可以预料，保持了足以发生可见固定数量

那么长的小群体指定比大群体的遗传力低。环境方差取决于培养或管理的条件:较大变化的条件降低遗传力;较一致的条件提高遗传力。最后,如果表型是两次或多次度量的平均值,根据度量的次数,遗传力将不同,也不同于一次度量的遗传力,其理由在第 8 章有关重复力部分已有解释。因此,无论何时对一特定性状的遗传力表述一个数值时,它必须被理解为是指特定条件下的一个特定群体的遗传力。在其他条件下其他群体中得出的数值有多少相同,要看群体结构和环境条件有多少相似而定。

对动物和植物中的许多性状已进行了相当多的遗传力测定。一些例子给在表10-1中。遗传力的估计不容易有很高的准确性,大多数估计值都有相当大的标准误。在同一生物体中对同一性状的不同估计值有大范围的变异,它们中的一些可能反映了被研究群体之间或条件之间的真实差异。尽管如此,在抽样误差的范围内,估计值在不同的群体中趋向于相似的。因为抽样误差大,所以,表 10-1 中的估计值被给定为最近于 5%。尽管缺少准确性,但很清楚,遗传力在性状间有很大的不同。而且,遗传力的大小和性状的属性之间有一定的联系。这在表 10-1 中可以看出。总的来说,具有最低遗传力的性状是与繁殖适合度最紧密相关的那些性状,而具有最高遗传力的性状则是那么一些性状,按生物学基础判断,作为自然适合度的决定因子,它们是最不重要的性状。对果蝇(Roff 和 Mousseau,1987)和大量物种的野生群体(Mousseau 和 Roff,1987)中不同性状遗传力的广泛调查,已充分证实了这种关系。为什么不同种类的性状应具有不同的遗传力,理由将在第 20 章中陈述。

表 10-1 各种动物中各种性状遗传力的近似值

	h^2(%)	来 源		h^2(%)	来 源
人			家禽		
体高	65	(1)	体重(32 周龄)	55	(7)
血清免疫球蛋白(IgG)水平	45	(2)	蛋重(32 周龄)	50	(7)
牛			产蛋量(到 72 周龄)	10	(7)
体重(成年)	65	(3)	小鼠		
乳脂(%)	40	(4)	尾长(6 周龄)	40	(8)
产奶量	35	(4)	体重(6 周龄)	35	(8)
猪			窝产仔数(第 1 窝)	20	(9)
背膘厚	70	(5)	黑腹果蝇		
饲料转化率	50	(5)	腹部刚毛数	50	(10)
日增重	40	(5)	身体大小	40	(11)
窝产仔数	5	(6)	卵巢大小	30	(12)
			产卵量	20	(11)

注:估计值四舍五入到近 5%;它们的标准误从大约 2%变化到大约 10%。

(1)西非群体。Roberts,Billewicz 和 McGregor(1978)。

(2)美国白人,Grundbacher(1974)。

(3)肉牛;很多估计值的平均。Preston 和 Willis(1970)。

(4)英国弗里生牛,第 1 泌乳期。Barker 和 Robertson(1966)。

(5)英国大白猪。Smith,King 和 Gilbert(1962)。

(6)英国大白猪。Strang 和 Smith(1979)。

(7)白来航品系杂种。Emsley,Dickerson 和 Kashyap(1977)。

(8)Rutledge,Eisen 和 Legates(1973)。

(9)Falconer(1965*b*)。

(10)Clayton,Morris 和 Robertson(1957)。

(11)Robertson(1957*b*)。

(12)Robertson(1957*a*)。

将遗传力的概念应用于植物，需要小心一些。当遗传力定义为 $h^2 = V_A/V_P$ 时，方差是单个值的方差。植物的单株值，特别是它们的产量，经常是没有效的，或者，如果有效，也会因竞争变得毫无意义，这种竞争在前一章中作为一种环境因子而被提及。产量通常以植物生长的小区的单位面积来表示。因此度量的单位是小区的产量，而不是单株产量。如果小区中的个体是一个家系——全同胞或半同胞的成员，“遗传力”就是家系间差异的遗传力，它的意义将在第 13 章中解释。本章的剩余部分仅涉及单个值的遗传力。

遗传力的估计

遗传力可从亲属间相似程度来估计。表 10-2 再次显示了前一章列出的表型协方差的组成。右边栏给出了以遗传力表示的回归或相关，从中可看出具有的某种关系

$$h^2 = b/r \text{ 或 } t/r \qquad [10.5]$$

其中 r 是协方差中加性方差的系数，或“理论”相关。因此，当以亲属间相关(或回归)的形式表示时，如果性状是完全遗传的，即如果所有的方差都是加性遗传的，那么遗传力就是所观察到的相关，与将发现的相关呈比例。

表 10-2　不同亲属类型的协方差

亲　属	协方差*	回归(b)或相关(t)
后代与单亲	$\frac{1}{2}V_A$	$b = \frac{1}{2}h^2$
后代与中亲	$\frac{1}{2}V_A$	$b = h^2$
半同胞	$\frac{1}{4}V_A$	$t = \frac{1}{4}h^2$
全同胞	$\frac{1}{2}V_A + \frac{1}{4}V_D + V_{Ec}$	$t \geq \frac{1}{2}h^2$

*上位互作的贡献忽略了，除全同胞外，对亲属可能存在的环境贡献也忽略了。

选择什么种类的亲属用于遗传力的估计取决于环境条件。除了哪一类的亲属实际上有用这一现实问题之外，还有两点需要考虑：精确度和偏差。一般而言，关系越近，估计值就越精确。其理由是对较远的亲属，观察到的回归或相关必须乘以一个较大因子($1/r$)，同时回归或相关的标准误必须乘以相同的因子以给出估计遗传力的标准误。统计的精确度在本章的后面部分将更全面地考虑。与精确度相比，通常对遗传力估计偏差的考虑更为重要。它由协方差的环境来源引起，在全同胞的情形下，还由显性所引起。从性状的生物学和实验设计考虑，我们必须确定哪一种协方差最不可能被环境组分所增大，这个问题已在上一章中讨论。一般说来，效应有时易于产生太高的估计值，例如大多数哺乳动物的身体大小就是这样(例 10.3 示证了由母体效应引起的偏差)。全同胞相关，它是协方差的环境组分显示在表中唯一的亲缘关系，是所有亲缘关系中最不可靠的。由共同环境引起的组分经常以大数量出现，而且很难通过实验设计克服；全同胞协方差被显性方差进一步增大了。因此全同胞相关很少使用，只不过为遗传力设置一个上限。

例 10.1

用 3 种不同方法确定了黑腹果蝇(*Drosophila melanogaster*)同一群体的腹部刚毛数的遗传力(Clayton, Morris 和 Robertson, 1957),有下列结果:

表 10-3 不同方法估计刚毛数的遗传力

估计方法	遗传力
后代—亲本回归	0.51 ± 0.07
半同胞相关	0.48 ± 0.11
全同胞相关	0.53 ± 0.07
联合估计值	0.52

用 3 种方法获得的估计值处于很满意的一致。在这种情形下,性状——刚毛数摆脱了由母体效应和共同环境引起的复杂性(表 10-3)。

现在让我们简要考察有关观察数据换算成遗传力估计值的一些技术问题。目前假定所有观察值是在一个没有选择亲本的随机交配群体中获得的。以后,将描述选型交配和选择的影响。

亲子回归

从后代对亲本的回归中进行遗传力的估计是比较容易做的。数据是以亲本的度量值——一个亲本或两个亲本的平均值——和它们后代平均值的形式获得的。于是从配对值的交叉乘积中计算出协方差。下面的例子说明对中亲值的回归。

例 10.2

图 10-1 示例了黑腹果蝇(*Drosophila melanogaster*)翅长的后代对中亲值的回归(Reeve 和

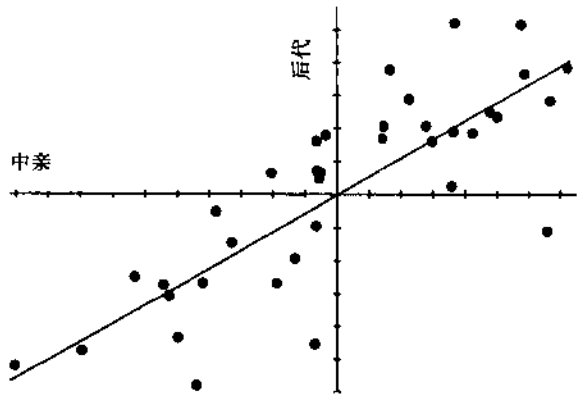


图 10-1 果蝇翅长的后代对中亲值的回归,这在例 10.2 中已有解释。中亲值沿着水平轴显示,后代的平均值沿着垂直轴显示(从 E.C.R. Reeve 博士热情提供的数据中绘制)

Robertson, 1953)。对 37 对亲本,以及每一对亲本平均 2.73 个后代进行了度量。亲本进行了选型交配,其结果是中亲值的方差比随机交配所得的方差要大,这一点将在较后的章节解释。图中每一点代表一对亲本的平均值(沿着水平轴度量)和它们后代的平均值(沿着垂直轴度量)。轴以 1 /100mm 的间隔作标记,它们在所有亲本和所有后代的平均值处相交。斜线是后代对中亲的线性回归。这条直线的斜率估计了遗传力,其值(± 标准误)为:$h^2 = b_{\overline{OP}} = 0.58 \pm 0.07$。

如果方差在两个性别中不相等，使用后代对亲本的回归会产生一种复杂性。在前一章中注意到，后代和中亲值的协方差等于加性遗传方差，条件是两性的表型方差相等。如果方差不等，严格地说，对中亲值的回归不能使用，并且必须对每一个性别单独计算遗传力。例如，雄性的遗传力，可从儿子对父亲或女儿对父亲的回归中来估计。可是，女儿对父亲的回归必须校正变异的差别，将雄性与雌性的表型标准差的比率与回归系数相乘。因此，如果 b 是女儿对父亲的回归，那么校正的回归是 $b' = b\,\sigma_{♂}/\sigma_{♀}$。类似地，雌性的遗传力可从女儿对母亲，以及通过 $\sigma_{♀}/\sigma_{♂}$ 校正儿子对母亲的回归来估计。来自 4 种回归的估计以及对不相等方差的校正，示例在下面的例子中。

例 10.3

在小鼠的一个随机交配品系中，用亲子回归估计了 6 周龄体重的遗传力（Falconer，1973）。雄性和雌性方差不相等，因此对后代和亲本的每一个性别单独计算了回归。表型标准差和它们的比率如下：

$$\sigma_{♂} = 3.786,\ \sigma_{♀} = 2.675,$$

$$\sigma_{♂}/\sigma_{♀} = 1.415,\ \sigma_{♀}/\sigma_{♂} = 0.707$$

表 10-4 给出了回归系数和它们的标准误，以及为校正方差的差异，两个回归系数必须乘以的因子。回归是所有后代对一个亲本的，因此回归和它们的标准误必须乘以 2 才可获得表 10-5 中给出的遗传力。根据后代的性别，估计值没有显著的不同，因此雄性和雌性后代平均起来列于表 10-5 的第三行。然而，雄性和雌性亲本之间估计值确实显著不同。雌性亲本极高的估计值可归因于母体效应引起的偏差。

表 10-4　具有标准误的回归系数和校正因子

	亲　本	
后代	雄　性	雌　性
雄性	0.110 ± 0.040	$(0.324 \pm 0.064) \times 0.707$ $= 0.229 \pm 0.045$
雄性	$(0.111 \pm 0.029) \times 1.415$ $= 0.157 \pm 0.041$	0.237 ± 0.043

表 10-5　遗传力及其标准误（%）

	亲　本	
后代	雄　性	雌　性
雄性	22 ± 8	46 ± 9
雌性	31 ± 8	47 ± 9
两性	27 ± 6	47 ± 6

同胞分析

从半同胞进行遗传力估计比初看起来更要复杂，需要更详细的注释。获得家畜数据的一个普通形式如下。许多雄性（公畜）每一个都与几个雌性（母畜）交配，雄性和雌性随机选择和随机交配。对每一个雌性的许多后代进行度量以提供数据。于是，所度量的个体形成一个半同胞和全同胞家系的群体。然后进行方差分析，通过它将表型方差分剖成不同成因的观察组分：由不同公畜后裔之间差异引起的方差组分（公畜间组分，σ_S^2）；由与同一公畜交配的不同雌性的后裔之间差异引起的方差组分（公畜内母畜间组分，σ_D^2）和由同一母畜后裔个体间差异引起的方差组分（后裔内组分，σ_W^2）。分析的形式显示在表 10-6 中。这里假设

有 s 头公畜，每头公畜与 d 头母畜交配，每头母畜产生 k 个后裔。均方值用 MS_S、MS_D 和 MS_W 表示。后裔内均方就是后裔内方差组分 σ^2_W 的本身估计值；但其余均方不是方差组分。指明方差观察组分的均方的组成显示在表的右边栏中，对它进行考察将看到方差组分是如何被估计的。例如，母畜间均方由后裔内组分和 k 倍的母畜间组分构成；因此母畜间组分被估计为：$\sigma^2_D = (1/k)(MS_D - MS_W)$。类似地，公畜间组分被估计为 $\sigma^2_S = (1/dk)(MS_S - MS_D)$，其中 dk 是每头公畜的后裔数。如果母畜的后裔数或公畜组中母畜数不相等，只要数目的差异不是很大，使用 k 和 d 的平均值不会有什么误差。准确的解决方法，这里描述起来太复杂，可参见 Snedecor 和 Cochran(1967，10.19 节)，Turner 和 Young(1969)，或 Searle(1971)。

表 10-6 半同胞和全同胞家系分析的形式

来 源	自由度	均方	均方的组成
公畜间	$s-1$	MS_S	$=\sigma^2_W + k\sigma^2_D + dk\sigma^2_S$
母畜间(公畜内)	$s(d-1)$	MS_D	$=\sigma^2_W + k\sigma^2_D$
后裔内	$sd(k-1)$	MS_W	$=\sigma^2_W$

s = 公畜数
d = 每头公畜所配的母畜数
k = 每头母畜的后裔数

下一步是要推断已从数据中估计出的观察组分和有因组分之间的联系，尤其是与加性遗传方差的联系，对它的估计是分析的主要目的。虽然所需的所有信息已经给出，可是对表 10-7 给出的观察组分的理解，也并不是不经解释就能立刻领会的。注意的第一点是表型方差的估计值由 3 种观察组分之和(σ^2_T)给出：$V_P = \sigma^2_T = \sigma^2_S + \sigma^2_D + \sigma^2_W$。这不一定等于从总平方和中估计出的观察方差，虽然两者很少有大的出入。现在考察公畜间组分 σ^2_S 的解释。这是半同胞家系平均值之间的方差，因此它估计了半同胞的表型协方差 $\mathrm{cov}_{(HS)}$，它是 $\frac{1}{4}V_A$。于是 $\sigma^2_S = \frac{1}{4}V_A$。接下来考察后裔内组分 σ^2_W。由于任何组间方差组分等于组内各成员的协方差，所以它表明组内组分就等于总方差减去组内各成员的协方差。母畜的后裔是全同胞家系，因此后裔内方差估计了 $V_P - \mathrm{cov}_{(FS)}$。这导致这样的理解：$\sigma^2_W = \frac{1}{2}V_A + \frac{3}{4}V_D + V_{Ew}$。最后，剩下了母畜间组分，而它的估计可通过下面的减法获得：

$$\sigma^2_D = \sigma^2_T - \sigma^2_S - \sigma^2_W = \mathrm{cov}_{(FS)} - \mathrm{cov}_{(HS)} = \frac{1}{4}V_A + \frac{1}{4}V_D + V_{Ec}$$

对公畜间和母畜间组分的考察，可以看出它们的和给出了全同胞协方差 $\mathrm{cov}_{(FS)}$ 的估计值，但这并没有为估计有因组分提供新的信息。关于方差的观察组分和有因组分之间联系的这些结论汇总在表 10-7。互作方差对观察组分的贡献可从对表 9-4 中给出的协方差的贡献中推导出来。

以下面两个例子阐明从同胞分析中进行遗传力的估计。对估计值标准误的计算是复杂的，Turner 和 Young(1969)对此作过描述。

表 10-7　同胞分析中方差观察组分的解释

观察组分	协方差和估计的有因组分	
公畜	$\sigma_S^2 = \text{cov}_{(HS)}$	$= \frac{1}{4} V_A$
母畜	$\sigma_D^2 = \text{cov}_{(FS)} - \text{cov}_{(HS)}$	$= \frac{1}{4} V_A + \frac{1}{4} V_D + V_{Ec}$
后裔	$\sigma_W^2 = V_P - \text{cov}_{(FS)}$	$= \frac{1}{2} V_A + \frac{3}{4} V_D + V_{Ew}$
总和	$\sigma_T^2 = \sigma_S^2 + \sigma_D^2 + \sigma_W^2 = V_P$	$= V_A + V_D + V_{Ec} + V_{Ew}$
公畜 + 母畜	$\sigma_S^2 + \sigma_D^2 = \text{cov}_{(FS)}$	$= \frac{1}{2} V_A + \frac{1}{4} V_D + V_{Ec}$

例 10.4

作为对从同胞分析中估计遗传力的说明，我们例举出以丹麦猪后裔测定站的记录为基础的丹麦长白猪的研究(Fredeen 和 Jonsson, 1957)。数据来自 468 头公猪，每头公猪配 2 头母猪，对来自每头母猪的 2 个雄性和 2 个雌性后裔的记录进行分析(表 10-8)。这里仅给出对雄性后裔体长的分析。表中显示的分析是在各测定站内和各年份内进行的，这就解释了自由度小于从上面所提数目中出现的适宜自由度的原因。从表 10-6 中显示的均方组成中计算了方差组分；这里 $k=2, d=2, kd=4$。在表底部显示的分析解释，通过忽略与这里无关的一些不重要的校正而稍为简化了。母畜间组分不大于公畜间组分，因此不可能有很大的非加性遗传方差和由共同环境引起的方差。因此，分别从公畜和母畜组分中获得的遗传力的两个估计值可认为是同等可靠的，而且它们以全同胞之间相似性为基础的联合可看作是最好的估计值(表 10-9)。

表 10-8　丹麦长白猪体长的同胞分析(数据仅来自雄性后裔)

来　源	自由度	均方	方差组分
公猪间	432	6.03	$\sigma_S^2 = \frac{1}{4}(6.03 - 3.81) = 0.555$
公猪内母猪间	468	3.81	$\sigma_D^2 = \frac{1}{2}(3.81 - 2.87) = 0.47$
后裔内	936	2.87	$\sigma_W^2 =$ 2.87
总　和			$\sigma_T^2 =$ 3.895

表 10-9　分析的解释

同胞相关	遗传力的估计值
半同胞：$t_{(HS)} = \frac{\sigma_S^2}{\sigma_T^2} = 0.142$	公畜组分：$h^2 = \frac{4\sigma_S^2}{\sigma_T^2} = 0.57$
	母畜组分：$h^2 = \frac{4\sigma_D^2}{\sigma_T^2} = 0.48$
全同胞：$t_{(FS)} = \frac{\sigma_S^2 + \sigma_D^2}{\sigma_T^2} = 0.263$	公畜 + 母畜：$h^2 = \frac{2(\sigma_S^2 + \sigma_D^2)}{\sigma_T^2} = 0.53$

例 10.5

我们还没有一个例子来说明共同环境在增大全同胞相关中的影响。这由小鼠的体重提供。表 10-10 给出的分析指出 6 周龄雌性小鼠的体重(J.C.Bowman,未发表)。有 719 个后裔来自 74 只公鼠和 192 只母鼠,每一母鼠产一窝。这些小鼠延续了 4 个世代,分析是在世代内进行的。由于每只母鼠的后裔数不同以及每只公鼠所配的母鼠数不相等,因此分析是复杂的。对这些不等数量进行的校正显示在用来估计组分的均方组成中,对此不必做解释。母畜组分比公畜组分大得多,表明共同环境或显性引起了实际的偏差。因此只有公畜组分能用来估计遗传力。获得的估计值是 $h^2 = 4 \times (0.48/5.14) = 0.37$(这个估计值的标准误为 0.26,这样它与 0 没有显著差异。这个实验的规模太小以致于没有多大实际用处,虽然它可用来说明这种方法)。根据表 10-7 给出的解释,现在可以从分析中来估计有因组分。区分共同环境和显性作为母畜和公畜组分间差异的原因是不可能的。表 10-11 中的估计值是基于这样的假设:即差异都是由共同环境引起的,从而 $V_D = 0$。在分析的解释中,我们可以比此更进一步,对显性方差规定一上限,结果如下。V_D 的最大可能值由后裔内组分 σ^2_W 设定。虽然很不像,但 $V_{Ew} = 0$ 是可能的,这将得出 $\sigma^2_W - 2\sigma^2_D = \frac{3}{4}V_D$,从中可得出 $V_D = 1.64 = 32\%$ 作为上限。于是 V_{Ec} 是 $\sigma^2_D - \sigma^2_S - \frac{1}{4}V_D = 1.58 = 31\%$ 作为下限。可是,有因组分的真实值很可能更接近表(ii)中的那些值。

表 10-10 小鼠体重的方差组成

来源	自由度	均方	均方组成	组 分
公畜	70	17.10	$\sigma^2_W + k'\sigma^2_D + dk'\sigma^2_S$	$\sigma^2_S = 0.48$
母畜	118	10.79	$\sigma^2_W + k\ \sigma^2_D$	$\sigma^2_D = 2.47$
后裔	527	2.19	σ^2_W	$\sigma^2_W = 2.19$
		$k = 3.48$;	$k' = 4.16$; $d = 2.33$	$\sigma^2_T = 5.14$

表 10-11 小鼠体重方差估计值

$V_P = \sigma^2_T$	$= 5.14 = 100\%$
$V_A = 4\ \sigma^2_S$	$= 1.92 = 37\%$
$V_{Ec} = \sigma^2_D - \sigma^2_S$	$= 1.99 = 39\%$
$V_{Ew} = \sigma^2_W - 2\ \sigma^2_S$	$= 1.23 = 24\%$

后代对母畜的公畜内回归

在具有上节所描述结构的群体里，通过在公畜组内计算的后代对母畜的回归，可以从亲子关系中估计遗传力。也就是说，对与一头公畜交配的每一组母畜单独计算后代对母畜的回归，并且来自每一组的回归合并成一个加权均值。这种回归估计了遗传力的一半，就像忽略了公畜的全面回归一样。公畜内回归比全面回归更好，因为它消除了有关回归直线上由公畜或公畜组之间环境差异引起的不想要的变异。Lush（1940）解释说，如果公畜组位于不同的畜群中，但母畜和每一公畜组的后裔饲养在同一畜群中，那么，该方法的主要优点是消除了由此产生的环境协方差，当然它还有其他优点。

联合估计值

对上面描述的典型实验设计，方差的有因组分是亲属间观察的协方差的简单线性函数。通过最小二乘统计分析（即回归、方差分析）对联系观察方差与其期望值的方程组求解，可获得方差组分估计值。从统计考虑，理想上这些设计应当是平衡的，对基于同胞协

方差的遗传力无偏估计，亲本必须从群体中随机取样（见下面）。与典型设计之一能提供的信息相比，经常有更多的信息是可用的。例如，当知道两个亲本及其大量后裔的表型时，采用上面描述的方法给出遗传力的两个估计值；一个来自后裔对中亲值的回归，另一个来自全同胞分析。类似地，基于经过几个世代的更遥远亲属获得信息也可用，不过采用标准设计不能从这些信息中估计出方差组分和遗传力，或者因仅提取了适于分析的部分数据而忽略了这类信息。而且，用于动物育种改良方案中的个体通常是经选择的群体，因此通过它们后裔方差分析获得的方差组分和遗传力估值是有偏的。

最大似然（ML）统计方法能适应数据中遗传关系的任何结构，因而利用了经适当加权的所有有效的信息；它们不需要平衡设计，并能考虑到亲本的选择。本质上，给定观察到的关系和要估计的参数（即方差组分、遗传力），ML 便评估观察数据能得到已给出的参数特定数值的概率。选择初值并使用迭代计算机算法来寻找似然值最大时的参数组合；这些是最大似然估计值（MLEs）。可是由于两个原因，MLE 是有偏的。首先，估计值被限定位于参数的范围内，因此方差组分不可能小于 0。其次，ML 假设固定效应是已知的，不考虑从数据中估计固定效应引起的自由度损失。在最简单的情形中，只有一个固定效应，即全群平均值，但在较复杂的数据组中，可以有很多固定效应；例如，奶牛数据中的场、年、季。这以同样方式导致了方差组分的低估计值，即采用 $\sum (X - \overline{X})^2/N$ 低估了从一个样本中估计的方差。采用修改的 ML 即约束 ML（REML）能消除偏差的第二个来源，REML 对观察值校正了固定效应的估计值。REML 分析在计算上是详尽复杂的，但现在它是估计遗传力和方差组分的首选方法。诸如 SAS 等通用软件程序包以及使用 REML 的专用遗传分析软件（例如，Shaw，1987；Meyer，1989*a*）都是有效的。ML 和 REML 分析的额外细节参见 Kennedy（1981）、Thompson（1982）、Shaw（1987）和 Meyer（1989*b*）。

双胞胎和人类数据

对人和牛来说，乍看起来同卵双生提供了估计基因型方差的一种方法。就像为实验动物或为植物提供了近交系，或者系间杂种一样，它们提供了基因型完全相同的个体。对人类双胞胎进行了很多研究，这些研究表明成对的成员在大多数性状上极其相似，甚至当他们从童年起就分开抚养。对牛双胞胎的研究，虽然规模小得多，却也显示了相同的情形（见 Hancock，1954；Brumby，1958）。采用它们表而的值，这些研究似乎表明对很多性状来说，遗传决定有很高的程度，达到 90% 或者甚至更多。可是，同卵双生在这方面的运用会因为共同环境引起的额外相似性而失效。双胞胎从受孕到出生直至它们在一起被抚养的期间都享有一个共同环境，因此成对之间的方差包含了混杂在遗传方差 V_G 内的共同环境引起的方差 V_{Ec}。通过比较两类双生子，同卵或单卵双生子（MZ）和异卵或二卵双生子（DZ），可部分地克服这种困难。二卵双生子是全同胞，和单卵双生子享有程度大约相同的共同环境。为了估计遗传方差的大小，我们问 DZ 双生子与 MZ 双生于有多少不相像。表 10-12 显示了双胞胎间和双胞胎内方差组分的组成，假定环境方差的两种组分 V_{Ec} 和 V_{Ew} 在 MZ 和 DZ 双生子中都相同。互作方差的贡献，为简单起见被忽略了，但能从表 9-4 中添加上。在两种组分上 MZ 和 DZ 双生于之间的差异估计了加性方差的一半以及显性方差的 3/4。可以注意到，对间均方的差异与组分之间的差异具有相同的期望值，即（MS_{MZ} –

MS_{DZ}）＝（$\sigma^2_{b(MZ)} - \sigma^2_{b(DZ)}$）＝$\frac{1}{2}V_A + \frac{3}{4}V_D$。同对双生子之间的相关是对间组分除以表型方差，因此 MZ 相关和 DZ 相关之间差异的二倍估计了（$V_A + \frac{3}{2}V_D$）/V_P。与遗传力（狭义）相比，这更接近于遗传决定程度（广义遗传力），可能这就是我们想从人类数据中获得的。可是应当注意到，即使假设环境组分相等，双生子分析并不提供 V_A/V_P 以及 V_G/V_P 的严格有效的估计值。如果上位组分加到协方差中，将会看到偏差会增加。例 10.6 说明了应用于人类 4 个性状上的双生子分析。

表 10-12 成对的双生子之间和之内的方差组分的组成（忽略了互作组分）

	对间，σ^2_b	对内，σ^2_w
同卵双生（MZ）	$V_A + V_D + V_{Ec}$	V_{Ew}
异卵双生（DZ）	$\frac{1}{2}V_A + \frac{1}{4}V_D + V_{Ec}$	$\frac{1}{2}V_A + \frac{3}{4}V_D + V_{Ew}$
差异	$\frac{1}{2}V_A + \frac{3}{4}V_D$	$\frac{1}{2}V_A + \frac{3}{4}V_D$

上面概述的双生子数据的分析关键取决于两个假设。第一，已经提及过，是方差的环境组分在两种类型的双生子中相同。第二，还未提及，是两类双生子的总遗传方差相同。而且，分析的目的是估计群体的参数，群体的大多数不是双生子，因此为了使这些估计值有效，双生子方差的环境组分必须与单个出生的个体环境组分相同。环境组分的差异，有很多可能的原因，下面是其中的一些原因（Stern，1973，更全面地解释和讨论了这些原因）。

（1）基因型—环境互作：正如第 8 章所解释的，在形式上这包括在环境方差中。它对 MZ 和 DZ 环境组分的贡献不相同。

（2）根据胎膜的排列，MZ 双生子有 3 种类型——单羊膜和单绒毛膜、单绒毛膜、或独立的羊膜和绒毛膜；所有的 DZ 双生子都是最后一种类型。

（3）子宫内同对双生子之间的竞争，这种竞争在 MZ 对子中比在 DZ 对子中可能更激烈。

（4）与独生子女相反，双生子是精确的同一世代。

（5）对双生子亲代的处理，这可能增强或消除相似性，可能不同程度地影响 MZ 和 DZ 双生子。

（6）对接合性（zygosity）诊断的失误。

（7）DZ 双生子间不同性别对的混入。

MZ 和 DZ 双生子总遗传方差之间的差异可以下面的方式产生（Nance，1976）。DZ 双生子的频率受包括种族差异在内的遗传因素的影响，而 MZ 双生子的频率，很少或根本不受遗传或种族因素的影响。因此，MZ 和 DZ 双生子样本可能不同地代表了群体的不同部分或阶层，因而遗传方差可能不同。虽然不能排除遗传和环境组分的平衡差异，但通过比较按 $\sigma^2_b + \sigma^2_w$ 估计的总方差，可以检验对方差相等的要求。如果总方差证明是相等的，并且对间或对内的比较没有明显的偏爱，则两种资料的信息可通过平均它们而联合起来

(Christian 等，1974)，即 $(\sigma^2_{b\mathrm{MZ}}-\sigma^2_{b\mathrm{DZ}})+(\sigma^2_{w\mathrm{DZ}}-\sigma^2_{w\mathrm{MZ}})=V_A+\frac{3}{2}V_D$。为了估计遗传力，$V_A+\frac{3}{2}V_D$ 的获得值要除以总方差 V_P。如果双生子的总方差与单生个体的不同，那么采用单生个体的 V_P 可获得适用于群体的遗传力。

尽管双生子分析有这么多困难，但由环境方差的不等性引起的偏差比全同胞相关的共同环境组分 V_{Ec} 引起的偏差可能要小。下面的例子说明了这一点。

例 10.6

表 10-13 给出了 MZ 双生子、同性别的 DZ 双生子和全同胞的 4 个性状的相关〔见 Huntley (1966)〕。这些性状全都在儿童时期度量，有十个手指上的总嵴数、按年龄校正的身高、语言智商测验以及社交成熟性得分，它是用来评估个体寻找其实际需要和承担与其年龄有关的责任能力的。选择这些性状，按表述顺序预测，将代表那些受环境影响会日趋加重的性状。从双生子差异中估计的遗传力，显示在表的底部，与这种预测是一致的。除嵴数之外，由于共同环境 V_{Ec} 引起的向上偏差，按加倍全同胞相关获得的估计值明显太高。双生子分析至少部分地消除了这种偏差。从亲子回归中估计的指嵴数遗传力大约为 0.8 (Mi 和 Rashad，1975)。单个手指数的遗传力较低，对不同的手指，遗传力从 0.58 变化到 0.68。总数的高值是由多次度量引起的，多次度量消除了所有环境组分，但单独影响每一个手指的 1/10 环境组分 V_{Es} 除外，其解释见第 8 章。

表 10-13　MZ 双生子、DZ 双生子和全同胞的手指嵴数、身高、智商和社交成熟性

	手指嵴数	身高	智商得分	社交成熟性得分
MZ 双生子	0.96	0.90	0.83	0.97
DZ 双生子	0.47	0.57	0.66	0.89
全同胞	0.51	0.50	0.58	0.32
$2(t_{\mathrm{MZ}}-t_{\mathrm{DZ}})$	0.98	0.66	0.34	0.16
$2t_{\mathrm{FS}}$	1.02	1.00	1.16	0.64

在解释人类亲属间相关，尤其是解释受文化传递影响的性状，共同环境的效应存在着严重的困难。这里不讨论在获得有意义的遗传力估计值的这些困难。只要这么说就足够了：通过利用几种不同种类亲属的相关以及拥有令每一个家系所受环境数量化的指数，可至少部分地克服这些困难（见 Morton，1974；Rao，Morton 和 Yee，1974，1976；Elston，1988 以及这方面的参考文献）。Eaves (1976) 以及 Cloninger，Rice 和 Reich (1979) 描述了估计文化传递效应的方法。例如，最后那些作者得出结论，在他们的数据中，智商得分的遗传力是 33%，但总的可传递方差是 69%。与试图估计诸如遗传力等遗传参数相比，检验参数是否非零可能更重要。这可通过“模型拟合”来进行。一个模型简单地就是以要检验的假设为基础的亲属间相关的一系列期望值。例如，假设可以是有共同环境引起的变异而没有遗传方差。如果数据对模型的期望给出显著坏的拟合，那么证明假设是错误的。Eaves 等 (1978) 综述了这些方法在人类心理性状中的应用。

选型交配

选型交配意味着相似的配相似的，它可从交配个体表型值之间的相关中看出。人类群体

的婚姻中，有很多数量性状是选型的，例如身高和智商得分等，虽然不一定被配偶慎重选择。在这一节中要考虑的问题是选型交配如何影响遗传力的估计，以及作为一种慎重的育种策略，应用选型交配是否在这方面有任何改进。选型交配的遗传后果是相当复杂的，这里仅给出结论，至多是简洁地表明它们是如何获得的。完整的解释由 Crow 和 Kimura(1970)给出。

选型交配的程度用交配个体表型值之间的相关 r 来表示，这是可观察到的。可是，遗传后果取决于配偶育种值之间的相关 m。为了推导 m 和 r 之间的联系，有必要知道什么支配着配偶的选择，相似性的主要原因是表型的、遗传的还是环境的。最初的表型相似性意味着配偶是以所研究性状的表型值为基础做选择的。这就是选型交配如何在育种方案中被应用。于是，可以看到两种相关之间的关系为 $m = rh^2$，其中 h^2 是配偶所选择性状的遗传力（这种关系的推导将在以后解释）。要描述的结果限制在最初的表型相似性作为选型交配的一种原因上。可是，人类的选型交配可能很少纯粹地由这种方式引起，将结果应用于人类数据需要谨慎，特别是在假设关系 $m = rh^2$ 可应用的情况下。

如果交配发生在遗传上或环境上互不相同的群体内，则会出现最初的遗传相似性或最初的环境相似性。在人类中出现的选型交配多少可能就是这样。于是观察到的表型相关 r 是由育种值或环境离差的“最初”相关产生的“次级”相关。最初相关不能从 r 中推导出来，除非它们中的一个可用其他方法估计出来，如果不知道 m，选型交配的遗传结果是不可能推导出来的。如果最初相关完全是环境的（$m = 0$），则将没有选型交配的遗传结果（除了增加中亲表型值方差会降低后代对中亲值的回归）。对于人类智商而言，环境相关可能是选型交配的基础。对智商得分的家族数据的分析看出，夫妇间表型相关($r = 0.5$)的大部分或完全可归因于“门当户对”的择偶标准（Rao，Morton 和 Yee，1976）。

回到根据表型值进行的选型交配上，遗传结果概述如下。育种值间引起的相关 m 导致了加性方差的增加，继而导致了遗传力的增加。可是，亲属间相关的增大比单独由遗传力增加导致的增加还要大。因此选型交配下遗传力的意义可能含糊不清。它可能被认为是亲属间相似性的决定，就像等式［10.5］所表达的，或是方差组分的比率 V_A/V_P，这两个数值在选型交配下是不相同的。这里将保留 V_A/V_P 的定义。我们主要关注的问题是：遗传力会增加多少，以及如何从亲属间相似性中估计遗传力（定义为 V_A/V_P）？

必须注意选型交配的其他方面如下：(1) 所有影响都不是即刻的；需要一些世代遵从随机交配才达到平衡状态。(2) 效应取决于影响性状的位点数；并假定数目足够大。(3) 对显性方差的影响很小，可被忽略。注意力将被限制在后裔中产生全同胞家系的成对交配上。不考虑连锁。

选型交配的结果可借助交配对子的协方差来获得，关于这一点需要三方面的知识。这些是给在表 10-14 中的两种相关 r(表型值之间)和 m(育种值之间)以及交配对子所属世代的方差组分。先前说明的关系 $m = rh^2$ 现在可获得，结果如下：$\mathrm{cov}(A_1A_2) = \mathrm{cov}(h^2P_1, h^2P_2) = h^4\mathrm{cov}(P_1P_2) = h^4rV_P = rh^2V_A$；根据表 10-14 中的(2)，$\mathrm{cov}(A_1A_2) = mV_A$，因此 $m = rh^2$。重要的是要注意到这里的 h^2 是在配偶选择发生的那个年龄上度量的性状遗传力。获得的后代方差如下。亲本育种值的协方差增加了中亲育种值的方差，因此增加了家系平均值之间的方差。家系内方差是由分离引起的，是不受影响的，只要分离的位点数不少。将家系间和家系内组分加在一起就给出了后裔世代的总方差。将加性和表型方差与随机繁殖基础群体的那些方差联系起来的等式给在表 10-14 中。在每一种情形下，第一个等式(4)和(6)指选型交配一个世代的后裔，

第二个等式(5)和(7)指已达到平衡的群体,此时方差保持稳定不变。加性方差除以表型方差就给出等式(8)和(9)中的遗传力。

表 10-14　选型交配。方差和协方差的近似表达式

配偶的协方差:

表型值, $\mathrm{cov}(P_1P_2) = rV_P$　(1)

育种值, $\mathrm{cov}(A_1A_2) = mV_A = mh^2V_P$　(2)

一个配偶的育种值与另一个配偶的环境离差,

$\mathrm{cov}(A_1E_2) = (r-m)V_A = (r-m)h^2V_P$　(3)

方差:	加性		表型		遗传力	
1 世代	$V_{A1} = V_{A0}(1+\frac{1}{2}m)$	(4)	$V_{P1} = V_{P0}(1+\frac{1}{2}mh^2)$	(6)	$h_1^2 = h_0^2\left[\frac{1+\frac{1}{2}m}{1+\frac{1}{2}mh^2}\right]$	(8)
平衡	$V_{A0} = V_A(1-m)$	(5)	$V_{P0} = V_P(1-mh^2)$	(7)	$h_0^2 = h^2\left[\frac{1-m}{1-mh^2}\right]$	(9)

亲属:	协方差		回归(b)或相关(t)	
后代,中亲	$\frac{1}{2}V_A(1+r)$	(10)	$b = h^2$	(13)
后代,单亲	$\frac{1}{2}V_A(1+r)$	(11)	$b = \frac{1}{2}h^2(1+r)$	(14)
全同胞	$\frac{1}{2}V_A(1+m)$	(12)	$t = \frac{1}{2}h^2(1+m)$	(15)

注释:

r = 配偶表型值之间的相关。

m = 配偶育种值之间的相关。当配偶的选择纯粹根据表型值时,$m = rh^2$。

h^2 = 交配年龄时的遗传力,定义为 V_A/V_P。

配偶的协方差:下标 1 和 2 指两个交配的个体;E 包括非加性遗传离差;V_A 和 V_P 指的是交配对子的世代。

方差:下标 0 指的是随机繁殖基础群体,1 指的是选型交配 1 世代的后裔。

亲属:显性 V_D 和共同环境 V_{Ec} 忽略了;V_A 和 h^2 指的是具有相关 r 和 m 的亲本世代。等式(10)到(14)适用于任一世代。等式(15)仅适用于一平衡群体;其他情形下,t = 等式(12)/等式(6) = 等式(15)/$(1+\frac{1}{2}mh^2)$。

借助这些等式,我们能够回答有关遗传力的两个问题。第一,选型交配一个世代可将遗传力增加多少?将 $m = rh^2$ 代入等式(8)看出它可将遗传力增加到为$(1+\frac{1}{2}h_0^2r)/\left(1+\frac{1}{2}h_0^4r\right)$的一个因子,其中 h_0^2 是起始遗传力。这种增加在提高从表型值预测个体育种值的准确性方面是有用的。可是,遗传力的增加决不是很大的,至多大约为 10%。采用这种方式在果蝇中运用选型交配的实验见 McBride 和 Robertson(1963)。第二个问题是:如果我们有一个处于平衡的群体,并且在该群体中估计遗传力,如果这个群体随机交配,那么遗传力将是多大?例如,将 $m = rh^2$ 代入等式(9)看出,如果在 $r = 0.5$ 的选型交配下遗传力为 0.50,那么随机交配的遗传力是 0.43。如果与其他物种进行比较,则这个问题与人类群体有关,但只有 m 已知或可猜测时,等式(9)才可适用于人类数据,因为对已经给出的理由来说,$m = rh^2$ 可能很少是正确的。

要考虑的最后一个问题是在选型交配的群体中定义为 V_A/V_P 的遗传力估计。考虑表 10-14 中(1)、(2)和(3)给出的亲代协方差,采用第 9 章描述的方法可以求得亲属的协方差。三种亲缘关系的协方差和回归或相关给在表 10-14 中(其他亲属的相关见 Nagylaki,1978)。这些协方差适用于任一世代,V_A 和 h^2 是亲代的值。正如等式(5)所示,V_A 的增加超过它的随机育种

值，协方差增加了一个因子，对后代和亲本该因子是$(1+r)$，对全同胞则为$(1+m)$。像以前一样，就单亲而论，亲子协方差与中亲值相同。而且，中亲值方差增加了与协方差相同的数量；因此和随机交配群体中的一样，后代对中亲值的回归就等于遗传力。这个结论对遗传力的估计有重要的实际影响。亲本之间的选型交配不影响后代对中亲值的回归，因此回归为亲本起源群体提供了有效的遗传力估计值。可是，选型交配在提高估计值精确性上有进步，因为中亲值方差增加致使标准误降低了，这一点将在下一节解释(细节参见 Reeve，1961)。

后代对单亲的回归和全同胞相关两者都受选型交配的影响。单亲方差简单地就是表型方差，但因为与未度量亲本的相关，后代对单亲的回归增加了一个因子 $(1+r)$，对完善的选型交配 $(r=1)$，它与对中亲值的回归是相同的。表 10-14 等式 (15) 中的全同胞相关忽略了显性和共同环境。如果假定可以忽略这些，那么通过代入 $m=rh^2$ 可以估计一个平衡群体的遗传力。于是方程是二次的，有 $h^2=[-1+\sqrt{(1+8rt)}]/(2r)$ 的解。

估计值的精确度和实验设计

由标准误表示的遗传力估计值的精确度很容易从用来估计遗传力的回归或相关的标准误中获得。令人不愉快，遗传力估计值的标准误很大，除非回归或相关是以很大数目为基础，因此重要的是做一切可能的事情来使标准误最小化。在计划估计遗传力的实验中，人们希望知道需要多少个观察值才能达到一给定的精确度；为了获得最大可能的精确度，在可用工具所强加的限制内，人们需要知道什么是最好的方法和最佳的实验设计。这些就是现在所要考虑的问题。方法是在后代对单亲或双亲的回归和同胞相关之间进行选择。可是对方法的选择，通常由多于准确度的实际考虑和消除偏差来决定。因此我们将不太注意方法的比较。设计问题涉及到每一个家系的个体数。能被度量的个体总数为空间、劳动力或费用所限。因此增加每一个家系的个体数目就减少了家系数。问题是要找到大家系和多个家系之间最好的折衷方案，以使得回归或相关的抽样方差最小。

在评价不同方法和设计的相对效率中，我们必须比较在同样规模上进行的实验；也就是说，在劳动力或费用上具有相同的总支出。因此我们必须首先确定限制实验规模的条件是什么。如果度量的劳动力是限制因素，例如像果蝇实验那样，那么限制就在于被度量个体的总数，如果亲本也要度量，则亲本数目也包括在内。另一方面，如果繁殖和饲养空间是限制因素，较大的家畜通常是这种情况，那么限制可能在于家系的数目或在于为度量而产生的后代总数，如果没有额外的费用，也可能包括亲本的度量。

这里我们不能考虑实验规模可能被限制的所有方面。为了说明的目的，限制将被认为是在一个世代中可被度量的个体数目，暗示相等数目的亲本和后代可被度量。Latter 和 Robertson (1960) 描述了找到亲子回归最优设计的原理，A.Robertson (1959a) 描述了找到同胞相关最优设计的原理。给出了结论，但对这些结论是如何获得的没有给出完整的解释。

亲子回归

假定 N 表示产生后代平均值和亲本平均值成对观察的家系数目；假定 k 表示每一个家系的亲本数目，它是 1 或 2；假定 n 表示每一个家系的后代数目。那么后代平均值对亲本平均值的回归系数的抽样方差可显示为：

$$\sigma_h^2 = \frac{k\,[1 + (n-1)\,t]}{nN} \quad (近似) \qquad [10.6]$$

抽样方差的这个近似表达式有利于人们比较度量一个或两个亲本的方法，以及决定每一个家系应当度量多少个后代。组内同胞相关 t 影响精确度，因为 t 低时与 t 高时相比，额外的后代提供更多的家系平均值信息。

一个亲本　首先考察只有一个亲本的度量。等式［10.6］中的分母 nN 是度量的后裔总数。如果这是限制实验规模的条件，那么 nN 被固定，当 $n=1$ 即 $(n-1)\,t=0$ 时抽样方差最小。因此在这些条件下的最有效设计是具有尽可能多的家系且每个家系只度量一个后代。于是，遗传力估计值的标准误如下：

$$标准误\ (h^2) = 2\ \sigma_h = 2/\sqrt{N} \quad (近似) \qquad [10.7]$$

为了达到 0.1 的标准误，有必要度量 400 个亲本和 400 个后代。这例证了这个事实：即使要得一个很适度的精确度的估计值，也需要很大的度量数目。如果只有 100 个家系可被度量，那么标准误将大约是 0.2，在大约 0.4 的情况下估计值将与 0 无显著不同。

两个亲本　现在对中亲值的回归考虑两个亲本的度量。如果每一家系只度量一个后代，将 $k=2$，$n=1$ 代入等式［10.6］中给出：

$$标准误\ (h^2) = \sigma_h = \sqrt{2/N} \quad (近似) \qquad [10.8]$$

可是，如果度量两个亲本，那么同样的环境将允许每一家系度量两个后代。将 $k=2$，$n=2$ 代入等式［10.6］给出估计值的标准误为$\sqrt{[(1+t)/N]}$，其中 t 是全同胞相关。比较将表明，在大多数情况下，对中亲值的回归比对一个亲本的回归可给出更好的精确度。

选型交配　亲本的选型交配提高了精确度。当然，必须度量两个亲本，因此只需要考虑对中亲值的回归。其效应来自中亲值方差的增加。后代的方差也增加了，但没有那么多，为简化起见这种增加被忽略了。选型交配下的中亲值方差是$\frac{1}{2}V_P\,(1+r)$，其中 r 是配偶之间的相关。其结果是选型交配的回归抽样方差大约是随机交配抽样方差的 $1/(1+r)$ 倍。因此，用标准误形式表示的精确度增加了一个因子$\sqrt{(1+r)}$，或如果选型交配是完全的，即如果 $r=1$ 时，则增加了一个因子$\sqrt{2}$。

大小不等的家系加权　在没有减少家系数的情况下，在每一个家系中度量尽可能多的后代经常是可能的。于是每一家系的后代数在家系间有所变化，这导致了如何根据后代数来加权家系的问题。适宜的权重取决于家系中后代之间的表型相关 t。加权的原理是大小为 n 的家系将按回归方差的倒数这一比例加权，回归方差是假定所有家系的大小都是 n 而取得的。Kempthorne 和 Tandon（1953）描述了这个加权。下面的方法（Falconer，1963）是校正了权重的一个修正式，以便大小 $n=1$ 的家系总会有权重 = 1。首先，组内相关 t 必须从后代的家系间和家系内的方差分析中计算出来。其次，要估计的回归系数必须能被推测，或近似地从未加权的家系平均值中估计出来。于是分配给 n 个后代平均值的权重 w_n 是：

$$w_n = (n + nB)/(1 + nB) \qquad [10.9]$$

其中对于在单个亲本的回归来说，$B=(t-b^2)/(1-t)$，对于在中亲值的回归来说，$B=(t-\frac{1}{2}b^2)/(1-t)$。权重对精确度没有很大的影响，除非 n 变化巨大。Bohren，McKean 和 Yamada（1961）检验了它的优点。

同胞分析

现在让我们考虑从全同胞或半同胞家系的组内相关中获得的估计值。为简单起见，开始我们假定半同胞家系不细分成全同胞家系，即在父系半同胞家系中只度量每个母畜的一个后代。设 N 是家系数，n 是每一家系的个体数，这样度量的个体总数为 $T = nN$。设 t 为组内相关。于是组内相关的抽样方差为：

$$\sigma_t^2 = \frac{2[1+(n-1)t]^2(1-t)^2}{n(n-1)(N-1)} \qquad [10.10]$$

当 $T = nN$ 的值为实验大小所限时，可以看出当近似地 $n = 1/t$ 时，组内相关的抽样方差最小。这样最适宜的家系大小取决于相关，因此取决于遗传力。假定由显性和共同环境引起的方差可以忽略不计，那么对全同胞 $t = h^2/2$，对半同胞 $t = h^2/4$。因此给出最有效设计的家系大小，对全同胞家系是 $n = 2/h^2$，对半同胞家系是 $n = 4/h^2$。在半同胞的情形下，我们假设每一母畜只度量一个后代，因此 n 就是每一公畜的最适母畜数。既然先前已知道遗传力至多只是近似的，因此预先不可能确切地确定最适宜的家系大小。可是，若与家系大小高出最适值相比，如果家系大小低于最适值，效率的损失要大得多。因此就误差而言有特大的家系结果会更好些。A.Robertson(1959a)指出，在事先缺乏对遗传力了解的情况下，半同胞分析一般应设计为其所用家系有 20～30 个个体。

当实验具有最优设计时，通过将 $n = 1/t$ 代入等式[10.10]可获得相关的抽样方差。求其近似值，这导致：

$$\sigma_t^2 = 8t/T \quad (\text{近似}) \qquad [10.11]$$

为了得到遗传力抽样方差，全同胞相关的方差必须乘以 4，半同胞相关的方差必须乘以 16。然后，将 $t = h^2/2$ 代入等式[10.11]，从全同胞中估计的遗传力抽样方差变为：

$$\sigma_{h^2}^2 = 4\sigma_t^2 = 16h^2/T \quad (\text{近似}) \qquad [10.12]$$

通过将 $t = h^2/4$ 代入等式[10.11]，来自半同胞的遗传力估计值的抽样方差变为：

$$\sigma_{h^2}^2 = 16\sigma_t^2 = 32h^2/T \quad (\text{近似}) \qquad [10.13]$$

因此，其他的情况都是一样的，根据它们的方差，全同胞家系的估计值的精确度是半同胞家系的两倍。

设计一个实验同时从亲子回归和同胞相关中来估计遗传力有时是理想的。Hill 和 Nicholas(1974)发现最优设计与两种方法单独运用时的最佳设计没有什么大的差异。

亲本选择

在实验群体和家畜中，所用的亲本经常是一个选择群。它们可根据要估计遗传力的性状，或根据与该性状相关的一些其他性状来选择。选择造成了亲本间方差下降，因而同胞的协方差也减少了。作为一种结果，从组内相关估计的遗传力偏低，比其真值要低 50%(见 Ponzoni 和 James,1978)。如果亲本的选择以要估计遗传力的性状为基础，那么它不影响亲子回归，无论是与单亲回归还是与中亲值回归，但是它降低了精确度，因为它减少了亲本方差(见 A.Robertson,1977a)。可是，如果两组亲本都选择，一组具有高值，一组具有低值，且只有来自这些选择群的后代被饲养，则选择能提高精确度。精确度的提高来自于将所有有用的方便条件专用于较极

端的家系，这为回归提供了最多的信息。当度量相等数目的后代和选择亲本时，每一组中选择亲本的最适宜比例大约是5%。细节可参见 Hill 和 Thompson(1977)。

习 题

10.1 假定由环境或显性引起的相似性可以忽略，从下列每一种相关或回归中估计的遗传力将是多大？

(1) 后代对父亲的回归	=0.21
(2) 后代对母亲的回归	=0.27
(3) 全同胞的相关	=0.34
(4) 后代对亲本均值的回归	=0.32
(5) 半同胞的相关	=0.02
(6) 雌性后裔对母亲姐妹的回归	=0.03
(7) 公畜内女儿对母畜的回归	=0.09

[解答序号：48]

10.2 下面的数据是在两个西非村庄成年居民的身高研究中获得的。女性身高校正到与男性相等，以便男性和女性的平均值相同。推导你认为最可靠的遗传力估计值是多大以及它的标准误。从这个数据中，除了 V_A 以外，还能获得协方差的什么其他组分？

	亲子回归±标准误		
	父亲	母亲	中亲
儿子	0.323±0.058	0.454±0.057	0.705±0.085
女儿	0.291±0.044	0.420±0.048	0.683±0.063
儿女	0.303±0.036	0.424±0.038	0.654±0.052
标准差（英寸）：男性 2.5；女性 2.3			

数据来自 Roberts，D.F. *et al*.（1978）*Ann. Hum. Genet.*，**42**，15～24。

[解答序号：58]

10.3 在习题 10.2 所使用的研究中，下面给出的同胞相关也获得了，人们的结婚习俗既提供了父系半同胞也提供了母系半同胞。性别在他们的相关中没有差异，因此结合在一起。运用这些数据来分剖这个群体的身高方差（组分的标准误在解答中没有给出）。这里估计的组分与习题 10.2 中估计的那些组分可做怎样的比较？

全同胞	0.406±0.035
父系半同胞	0.140±0.056
母系半同胞	0.257±0.101

[解答序号：68]

10.4 证明可以从方差分析中估计双胞胎的组内相关 $t=(B-W)/(B+W)$，其中 B 和 W 分别为对子间和对子内的均方。

[解答序号：78]

10.5 皮肤褶的厚度为人的肥胖提供了有用的度量。下表给出了 10 岁以下以及 10 岁和 15 岁之间双胞胎中皮肤褶厚度的相关。估计两组的遗传力。关于由共同环境引起的相似性，

可推断出什么？

数据来自 Brook，C.G.C.*et al*.（1975）*Brit*.*Med*.*J*.，**1975**，2，719～721。

	10岁以下	10～15岁
MZ	0.64	0.91
DZ	0.38	0.42

[解答序号：88]

10.6 下表给出了来自黑腹果蝇（*Drosophila melanogaster*）一随机繁殖群体的方差分析的均方。性状是腹侧刚毛数。这些是位于胸侧的刚毛，计数了一侧的刚毛数。有62只雄果蝇（公蝇），每一只配3只雌果蝇（母蝇）。每一只雌果蝇在一个小瓶中产卵。对每一只雌果蝇的10个雄性和10个雌性后代的刚毛进行计数。计算半同胞和全同胞的相关，并估计出可被这些数据区分开的方差组分。

	雄性	雌性
公蝇间	3.894	4.461
公蝇内母蝇间	2.198	2.061
母蝇内	1.125	0.893

数据来自 Sheridan，A.K.*et al*.（1968）*Theor*.*Appl*.*Genet*.，**38**，179～187。

[解答序号：98]

10.7 证明如果全同胞相关为0.75，那么该性状的遗传力不可能大于0.5。小鼠断奶体重的全同胞相关为0.8；与此相应的最大遗传力是多大？

[解答序号：108]

10.8 对位于Galapagos群岛的达尔文雀科鸣禽的一个种，*Geospiza fortis* 群体形态变异的一项研究提供了下列有关喙深的数据。你将如何解释这些数据？

回归±标准误

后代—中亲	0.82±0.15
后代—父亲	0.47±0.17
后代—母亲	0.48±0.13

相关±标准误

全同胞	0.71±0.12
父亲—母亲	0.33

数据来自 Boag，P.T.& Grant，P.R.（1978）*Nature*，**274**，793～794。

[解答序号：118]

10.9 假定一个群体经选型交配繁殖了足够长时间已达到平衡。配偶之间的相关 $r=0.4$，配偶的选择纯粹基于一特定性状的表型值。关于这个性状的全同胞相关为 $t=0.3$。已知性状有可以忽略的显性方差，且性状几乎不受共同环境对同胞的影响。在这个群体中，遗传力的估计值是多大，如果没有选型交配，则遗传力将是多大？

[解答序号：128]

10.10 如果下面给出的遗传力是用指明的方法估计的，且在每一种情形中度量的个体总数都为400，假定同胞之间没有环境相似性，且数据来自一随机繁殖群体的未选择的个体，那么估计值的标准误大约是多大？

（1）$h^2=0.5$，来自儿子对父亲的回归；200个父亲。

（2）$h^2=0.6$，来自3个后代（全同胞）的平均值对它们亲本平均值的回归。

（3）$h^2=0.4$，来自大小为5的家系中全同胞的相关。

（4）$h^2=0.2$，来自20个半同胞家系中的半同胞相关，它们之间没有全同胞。

[解答序号：138]

第 11 章　选择：Ⅰ．反应及其预测

至此，我们对数量性状的论述主要关注的是对随机交配下生存的群体遗传特性的描述，还没有考虑倾向改变其特性的那些影响；现在我们必须考虑由育种者或实验者行为所引起的变化。育种者行为有两种途径能改变群体的遗传特性：第一是通过对用作亲本的个体挑选，这形成了选择，第二是通过对亲本交配方式的控制，这包括近交和杂交。以一种形式或另一种形式进行的选择是家畜和植物获得所有改良的手段。因此，在这一章中，我们开始考察数量遗传学最重要的应用。选择意味着从“最好的”个体中育种，无论“最好的”是什么。在这方面，数量遗传学的理论可以帮助的方法是，首先，通过表明如何选择具有最好育种值的个体，其次，通过预测结果，以便能对不同的育种方案作比较。选择的最简单形式是根据它们自己的表型值来选择个体。这是在本章中要考察的选择形式。第 13 章将阐明来自亲属的信息如何有助于识别具有最好育种值的个体。在实验动物中进行的实验选择提供了研究数量性状遗传学的一种手段。选择的这个方面将在下一章中详述。

在任一实际的选择计划中，所用亲本的数目或多或少受到限制，其结果是一些近交不可避免地发生了，它的影响添加到选择的效应中。开始，都将忽略可能有的任何近交效应，但后来必须考虑近交效应。

选择的基本效应是以第 2 章中描述的方式改变基因频率的列阵。可是，基因频率本身的改变现在几乎完全是隐而不见的，因为我们无法研究与一个数量性状有关的单个位点。因此，可观察到的选择效应主要局限于群体平均值的改变。可是，让我们采用一般的术语来稍稍进一步地考察基因频率的潜在变化。

为了描述从一个世代到下一个世代遗传特性的变化，我们必须在个体生命周期的同一点上来比较连续的世代，并且这一点是由年龄固定的，就在这点上度量所研究的性状。最常见的是，大约在性成熟的年龄度量性状或在年轻的成年个体上度量性状。亲本的选择在度量之后进行，这些经选择的个体之间的基因频率不同于它们在选择前整个群体中的频率。如果被选择个体之间的生殖力以及它们后裔之间的存活力没有差异，那么后裔世代中的基因频率与选择亲本的相同。这样，人工选择，即在亲本的选择中由育种者行为引起的选择，通过将亲本世代的成年个体分成选择和淘汰的两组，这两组基因频率不同，由此产生了基因频率的变化。自然选择，通过亲本个体之间生殖力或其后代之间存活力的差异面起作用，可能造成了亲本个体和在后裔世代中被度量的个体之间基因频率的进一步变化。这样选择在 3 个阶段上可能引起基因频率的变化：第一是通过在亲代成年个体之间的人工选择；第二是通过在亲代成年个体之间生殖力的自然差异；第三是通过子代个体之间存活力的自然差异。虽然生殖力和存活力的自然差异总是存在的，但它们不一定总是相关的，因为它们不一定跟和数量性状有关的那些基因发生联系。

选择反应

我们主要感兴趣的由选择产生的变化是群体平均值的改变。这就是选择反应，用符号 R 表示；它是被选择亲本的后代与选择前整个亲代之间平均表型值的差异。所采用的选择度量是被选择亲本的平均优势，它被称为选择差，用符号 S 表示。它是选作亲本的个体平均表型值，以偏离群体平均值的离差表示，群体平均值是进行选择之前亲代中所有个体的平均表型值。为了推导反应和选择差之间的关系，让我们设想一个随机交配群体的两个连续世代，这以草图的形式显示在图 11-1 中。每一个点代表一对亲本和它们的后裔，根据沿水平轴度量的中亲值和沿垂直轴度量的后裔平均值来定位这些点。原点代表群体平均值，假定它在两代中相同。斜线是后裔对中亲值的回归直线（根据实际数据绘制的这类图解给在图 10-1 中）。现在让我们认为对亲代的一组个体做过选择的，并假定它们具有最高的值。这些成对的亲本和它们的后裔在图中用实心圆点表示。对亲本根据它们自己的表型值进行选择，而不管后裔或任何其他亲属的值。让 S 代表这些选择亲本的平均表型值，以偏离群体平均值的离差形式表示。类似地，用 R 表示它们的后裔偏离群体平均值的平均离差。于是 S 就是选择差，R 就是反应。由十字标记的点代表被选择亲本和它们后裔的平均值，它的预期位置在所显示的回归线上。这样，比率 R/S 等于回归线的斜率。因此反应和选择差之间的关系由下式给出：

$$R = b_{O\bar{P}}S \qquad [11.1]$$

在上一章我们看到后裔对中亲值的回归等于遗传力，只要后裔和亲本之间没有非遗传原因引起的相似性。对此，我们必须附加一个条件，即应当没有自然选择：也就是说，生殖力和存活力与所研究性状的表型值没有相关。因此，只要这些条件成立，那么反应与选择差的比率就等于遗传力，并且反应可由下式给出：

$$R = h^2S \qquad [11.2]$$

等式［11.2］表示的反应和选择差之间的关系直接来自遗传力的含义。在上一章（等式［10.2］），我们注意到遗传力等于个体育种值对其表型值的回归。根据定义，后裔与群体平均值的离差就是亲本的育种值，因此反应等于亲本的育种值。这样它表明后裔的期望值可由 $R = h^2S$ 给出。

在反应等式的推导中，设想的情形有一点与实际选择不符合。我们假定亲代个体随机交配，随后才实施选择。可是实际上，根据个体值而不是中亲值的选择在交配前就实施了。这种情况的影响是，那些被作为整个亲本群体的一部分个体进行了选型交配。然而，选型交配对中亲回归几乎没有影响，这一点在上一章我们注意到了，因此选择过程的这种特性可被忽略。

应当提及的另一点涉及回归的线性问题。在图 11-1 中它被绘制为一条直线并假定它在等式［11.1］中是线性的。在大多数情况下，这个假设是有理由的，等式［11.1］对一个大致近似是正确的。非线性可由显性产生。考虑图 7-2 将看到如果所有的方差都是由显性的一单个位点引起，那么育种值对基因型值的回归将是非线性的。可是，与少数位点相比当有更多的位点时，基因型值的分布几乎是正态的了，这一点可从图 6-1 中看出，回归实际上是线性的（Gimelfarb，1986）。非线性回归的讨论还可参见 Robertson（1977*b*）。

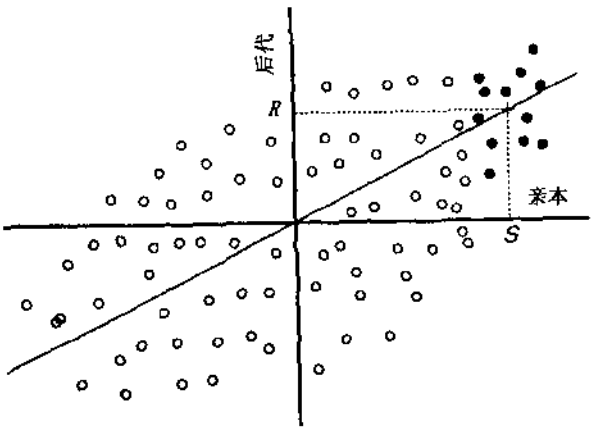


图 11-1　后裔平均值对中亲值绘制的图，以说明在本文中解释的选择反应

反应的预测

这些反应等式的主要用途是预测选择反应。让我们更进一步地考虑要进行的预测属性。首先，很显然，等式［11.1］不是一种预测而仅仅是一种描述，因为直到已经养育了后裔世代才能度量亲子回归。可是，根据从以前世代中获得的遗传力知识，等式 $R = h^2S$ 提供了预测的一种方法。预测中使用的遗传力可以通过任何方法例如同胞相关来估计，而不必从亲子回归来估计。直到亲本已选择之后才能知道选择差 S，但可以预测它的期望值，这一点将在下一节解释。下面的例子说明选择差和选择反应的计算，以及根据等式［11.2］进行的反应预测。

例 11.1

表 11-1 的数据来自 Clayton，Morris 和 Robertson（1957）对黑腹果蝇（*Drosophila melanogaster*）腹部刚毛数选择的实验。刚毛数的遗传力首先在选择前的基础群体中估计，得出其值为 0.52，这在例 10.1 中已表述了。选择高刚毛数的亲本具有一个平均优势为 $S = 40.6 - 35.3 = 5.3$ 根刚毛。根据等式［11.2］，预测的反应是 $0.52 \times 5.3 = 2.8$。观察到的反应是 $37.9 - 35.3 = 2.6$ 根刚毛。

表 11-1　黑腹果蝇腹部刚毛数的选择实验结果

世代	所有度量个体的均值	被选择的个体均值	选择差	反应	
				期望值	观察值
亲代	35.3	40.6	5.3	2.8	—
后代	37.9	—	—	—	2.6

从原理上说，反应的预测仅对一个世代的选择是有效的。反应取决于亲本被选择世代中性状的遗传力，因此严格地讲，不重新确定每一个世代的遗传力，以后世代的反应就不能被预测。为什么预测遗传力会改变，这有两个原因。首先，如果有反应，则基因频率一

定会改变，而遗传力取决于基因频率。这种变化相当长时间似乎不那么明显，因为基因频率的变化小，除非仅包含几个位点。其次，对亲本的选择减少了方差和遗传力。这发生在早期世代中。它将在以后作简要解释，同时予以忽略。可是，遗传力的这些预期变化不大，实验已经表明经过几个世代——达到 5、10 个世代或甚至更多的世代，反应通常保持，几乎没有变化。这一点将在本章后面和下一章中给出的选择反应图中看到。

选择差和选择强度

选择差可事先预测，只要两个条件成立：被选性状的表型值呈正态分布以及选择是截断选择。截断选择意味着严格按照由它们表型值判断的优势次序来选择个体，被选择的个体都优于任何被抛弃的个体。在这些条件下，选择差仅取决于选择群中包括的群体比例以及性状的表型标准差。选择差对这两个因素的依赖性在图 11-2 中以图解形式进行了说明。图显示了表型值的分布，并假定它是正态的。假定具有最高值的个体是要选择的，这样分布在截断点处被明显地分开，高于这个值的所有个体被选择，低于这个值的所有个体被淘汰。每张图中的箭头标出了选择组的平均值，S 是选择差。在图（a）中，一半的群体被选择，选择差是相当小的；在图（b）中，只有 20% 的群体被选择，选择差大得多。在图（c）中，也只有 20% 的群体被选择，但表示的性状很少有变异，因此选择差较小。（c）中标准差是（b）中标准差的一半，选择差也只有（b）的一半大。

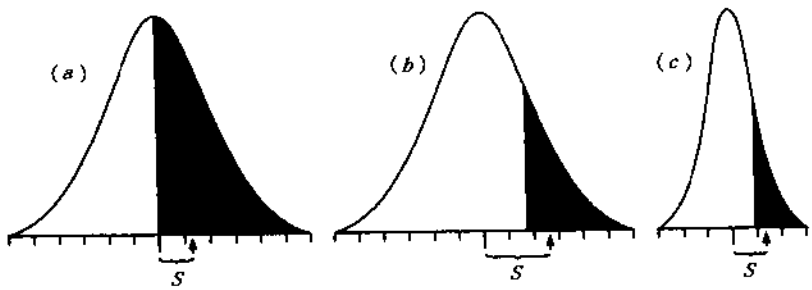


图 11-2 图用来显示选择差 S 如何取决于被选择群体的比例以及一个正态分布性状的变异性

处在阴影面积内的、超出截断点的所有个体都被选留。在各横坐标上都标有假想的度量单位。

（a）50% 被选留；标准差 2 个单位：$S = 1.6$ 个单位。

（b）20% 被选留；标准差 2 个单位：$S = 2.8$ 个单位。

（c）20% 被选留；标准差 1 个单位：$S = 1.4$ 个单位。

度量变异性的标准差是性状和群体的一种特性，它规定了表示反应的单位，即多少磅、毫米、刚毛等。如果选择差用表型标准差 σ_P 的形式表示，那么可使选择反应一般化。这种标准化的选择差 S/σ_P 被称为选择强度，用符号 i 表示。于是选择差是：

$$S = i\sigma_P$$

等式［11.2］中的期望反应变为：

$$R = ih^2\sigma_P \qquad [11.3]$$

注意到 $h = \sigma_A/\sigma_P$，其中 σ_A 是育种值的标准差（加性遗传方差的平方根），我们可以用如下形式书写这个等式：

$$R = ih\sigma_A \qquad [11.4]$$

这有时用于不同选择方法的比较。

选择强度 i 仅取决于选择群中包括的群体比例，并且，只要表型值的分布是正态的，它就可从正态分布特性表中被确定。如果 p 是选择的比例，即落在超过截断点的群体比例，z 是截断点处纵坐标高度，则从正态分布的数学特性中得出：

$$\frac{S}{\sigma_P} = i = \frac{z}{p} \qquad [11.5]$$

因此，只要给出选择的比例 p，我们就能找出选择个体的平均值将超出群体平均值多少个标准差：也就是说，选择强度 i。图 11-3 中的图解显示了 i 和 p 之间的关系。对于给定 p

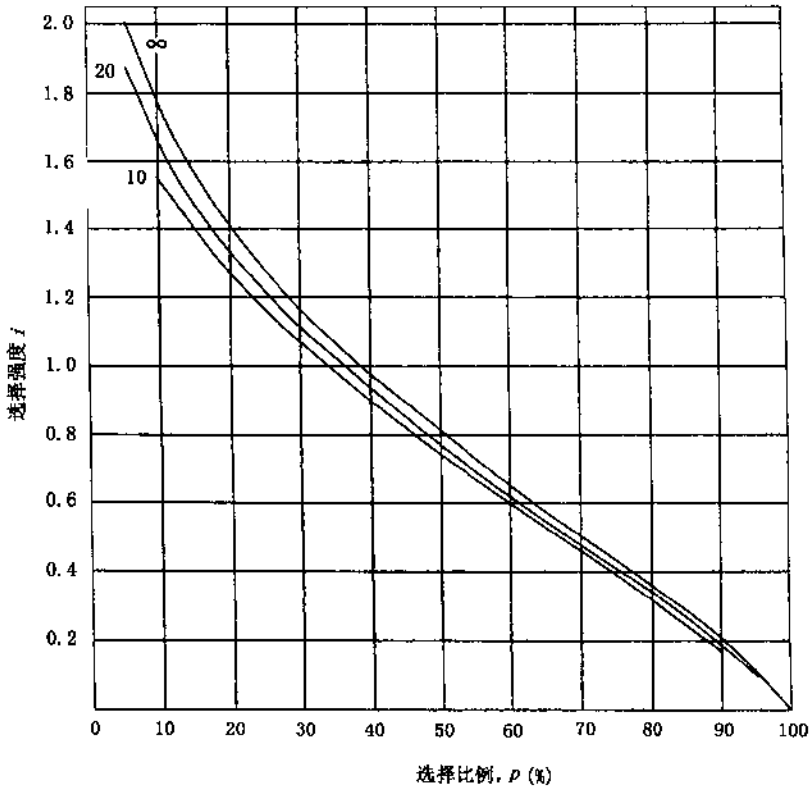


图 11-3　选择强度与选择比例的关系

选择强度是以表型标准差为单位的被选择个体的平均离差。上面的一条曲线指的是选择来自于被度量个体的总数足够大；下面两条曲线指的是选择分别来自于总数为 20 和 10 个个体的样本。假定其服从正态分布。

值的 i 值以表的形式列在附录表 A 中。严格地说，等式［11.5］中给出的 i 和 p 之间的关系仅适用于大样本：也就是说，当大量的个体被度量，并且选择是在它们之间进行时。当选择是在小数目度量个体中进行时，选择群的平均离差稍微小一些。选择强度可从排序数

据的离差表中找出（Fisher 和 Yates 的表 XX，1963）。图 11-3 中下面两条曲线显示了当选择在 20 和 10 个度量个体的样本中进行时的选择强度。附录表 B 给出了当选择是在小数目的样本中进行时的一些 i 值。

例 11.2

Clayton，Morris 和 Robertson（1957）通过研究果蝇（*Drosophila*）腹部刚毛数，进行了不同选择强度下期望和观察反应的比较。首先用 3 种方法确定遗传力，产生了一综合估计值为 0.52（见例 10.1）。刚毛数（两性平均）的标准差是 3.35。选择以 4 种不同强度进行了 5 个世代，包括上选和下选（即增加和减少刚毛数）。在每一种情形下，都选择 20 只雄性和 20 只雌性作为亲本，选择强度因被选择的亲本原始数目不同而变化，这一点显示在表11-2的第一栏。与这些选择比例相对应的选择强度给出在表的第二栏。于是从等式［11.3］中可得到期望反应。例如，在最强烈的选择下，它的 $R = 1.4 \times 3.35 \times 0.52 = 2.44$。观察反应给在表的最后两栏。虽然它们并不很精确地与期望值相符，但它们表明了因选择强度降低使选择下降而如何引起变化。

表 11-2 果蝇腹部刚毛数不同选择强度下的期望值与观察值

选择比例 p	选择强度 i	每世代的平均反应		
		期望值	观察值	
			上选	下选
20/100 = 0.20	1.40	2.44	2.62	1.48
20/75 = 0.267	1.23	2.14	2.20	1.26
20/50 = 0.40	0.97	1.65	1.46	0.79
20/25 = 0.80	0.34	0.59	0.28	-0.08

等式［11.2］中的选择差 S 和等式［11.3］中的选择强度 i 指所用全部亲本的平均优势。当应用雄性和雌性时，对其选择的数量可以不同。例如，一些性状仅在一个性别上能度量，因此对另一个性别便不能进行选择。如果对雄性和雌性所做的选择不同，则所采用的 S 或 i 值是两个性别的未加权平均值，即：

$$S = \frac{1}{2}(S_m + S_f) \qquad [11.6a]$$

或

$$i = \frac{1}{2}(i_m + i_f) \qquad [11.6b]$$

例如，如果只有雌性可被选择，那么 $S = \frac{1}{2}S_f$，期望反应为 $R = \frac{1}{2}h^2 S_f$。若注意到 $\frac{1}{2}h^2$ 是后裔对单个亲本的回归，那么这便能与等式［11.1］联系起来。用作亲本的性别数量也可能不同。于是，S 或 i 的值仍由等式［11.6］给出，因为后裔中的一半基因来自亲本的每一个性别，而不管其数量多少。

反应的改良

育种者用来改良反应速度的方法可从等式 $R = ih^2\ \sigma_P$ 中看出来。表型标准差 σ_P 仅仅

指定了度量的单位。通过注意饲养和管理技术来减少环境变异、当多次度量可能时通过多次度量、以及在小范围内通过上一章所解释的选型交配都能够提高遗传力。初看起来，增加选择强度似乎是改良反应的一种容易进行的方法，但有两个因素限制着育种者应用这种方法的作为。首先，生物体的繁殖速度限制了选择强度，因为用于育种的选择比例决不能低于更新所需的比例。也就是说，更新每一对亲本平均需要两个个体。因此生物体越多产，可运用的选择就越强烈。如果雄性和一头以上的雌性交配，则雄性比雌性具有更多的后裔，同时对雄性的选择比对雌性的强。例如，假定每一个雄性配 10 个雌性，每一个雌性平均有 5 个女儿。为了供雌性的更新用，对雌性的选择比例不能低于 1/5，但雄性平均有 50 个儿子，则允许选择 1/50 来更新雄性。在这种情形下，选择强度的上限对雌性将是 $i_f = 1.40$，对雄性则为 $i_m = 2.42$，由等式［11.6b］给出一净强度为 $i = 1.91$。限制选择强度的第二个因素是群体大小和近交的结果。近交几乎总是降低繁殖适合度以及与它相关的性状，这一点将在第 14 章中解释。因此当近交衰退与来自选择的增益相抵消时，用作亲本的数目必须大到足以保持这种近交衰退到一个可接受的水平。例如，在实验工作中，人们可能决定使用不少于 10 或 20 对的亲本。当用这种方式所使用的亲本数被固定时，只有通过对选作亲本的个体做更多数目的度量才能提高选择强度。

世代间隔　对于选择的有效后代数不仅取决于亲本的繁殖速度，而且在很多生物中取决于育种者进行选择之前他愿意等待多久。每单位时间的进展通常比迄今所叙述的每世代的进展更重要。因此在计算选择反应中世代间的时间间隔是一个重要的因素。育种者通过等待在他进行选择之前养育更多的后裔，便能够提高选择强度和每世代的反应，但这样做则不可避免地增加了世代间隔，由此可能降低了每单位时间的反应。因此在选择强度和世代间隔之间存在利害冲突，必须在这两者之间找到最佳的折衷方案。增加后代数将至一定量，超过此量将不再有益。后代的最适宜数目不能用一般的方式来表述，每一种情形必须根据它的特定环境条件来计算。

在计算任何选择方案下的世代间隔过程中，必须区分不连续的和重叠的世代。当世代不连续时，后代一直被饲养到最后出生的成熟为止；然后进行选择，中选个体在大约同一时间全都交配。世代间隔是连续世代中进行交配之间的间隔。当世代重叠时，由中选后代对亲本进行更新，或多或少是一个连续的过程，中选后代一经成熟即行交配。世代间隔可计算为亲本的中选后代出生时双亲的平均年龄。问题是要找到淘汰亲本的最适宜年龄。下面的例 11.3 说明世代不连续以及世代重叠时最适宜的计算方法。当使用的雄亲比雌亲少时，雄性和雌性的世代间隔 L_m 和 L_f 必须像选择强度 i_m 和 i_f 一样区分开来。比率 $(i_m + i_f)/(L_m + L_f)$不得不最大化后是多少。Ollivier（1974）解释了这个答案，能在给出主要家畜的答案图解中找到它。有时，根据雄性和雌性亲本是否被用来产生儿子或女儿，有必要区分 4 种选择强度和世代间隔。于是整个的比率 i/L 被计算为$\sum i/\sum L$，其中$\sum i$ 是 4 个强度的和，$\sum L$ 是 4 个世代间隔的和（Rendel 和 Robertson，1950）。

例 11.3

让我们假定对小鼠的某一性状进行选择，每单位时间进展的速度是目标。问题是：应当饲养多少窝？为了求出产生最大进展速度的窝数，我们必须求出选择强度和世代间隔。然后两者的比率将给出相对速度。通过乘以遗传力和标准差可以获得实际的速度，但这些

因素将被假定与饲养的窝数是独立的。对每周进展的期望速度的比较列在表 11-3。对两个不同的平均窝大小进行了比较意味着对每窝饲养的幼鼠进行了比较。假定被选择的性状在性成熟前可度量，并且是父母 9 周龄对第一窝出生了，紧接着以 4 周的间隔出生随后的一窝。也假定群体足够大，在计算选择强度中可作为一个大样本来处理，还假定选择了相等数目的雄性和雌性。

世代间隔取决于选择和交配的过程。考虑两种不同的过程。首先，选择被推迟直到所有的窝都被度量为止。然后所有中选的小鼠在同一时间交配，世代是不连续的。列在表中 L 项下的世代间隔是被饲养的最后一窝出生对父母的年龄。对实验室实验这是实际的方法。其次，所需要的小鼠从饲养的所有窝中等量地选择出来。例如，如果饲养了两窝，从第一窝中选择每窝一个，从第二窝中选择每窝一个，使得每一家系的总数为 2 个。如果总数很大，选择强度与第一种过程的相同。中选的小鼠一成熟就交配，因此世代是重叠的。列在表中 $\bar{L}$ 项下的世代间隔，是其窝仔鼠出生时父母的平均年龄。这种过程对实际育种方案更实际。

最适宜的窝数由比率 i/L 或 $i/\bar{L}$ 的最大值来表示。根据窝大小和过程，它有所变化。对第一种过程，如果每窝饲养 6 只幼鼠，则仅饲养一窝就可获得最大反应速度；如果饲养 4 只幼鼠，等待第二窝是值得的，但等待第三窝就不值得了。对于第二种过程，如果饲养 6 只幼鼠，则亲本在其第二窝之后就应淘汰；如果饲养 4 只幼鼠，当亲本在其第二窝或第三窝之后被淘汰对，进展的速度是相同的。这些结论如果没有表中显示的计算是很难猜得出来的。

表 11-3 单位时间进展的期望速度

N	L	$\bar{L}$	$n=6$				$n=4$			
			p	i	i/L	$i/\bar{L}$	p	i	i/L	$i/\bar{L}$
1	9	9	0.333	1.10	0.122	0.122	0.500	0.80	0.089	0.089
2	13	11	0.167	1.50	0.115	0.136	0.250	1.27	0.098	0.115
3	17	13	0.111	1.71	0.101	0.132	0.167	1.50	0.088	0.115
4	21	15	0.083	1.85	0.088	0.123	0.125	1.65	0.079	0.110

n = 每窝饲养的幼鼠数。

N = 饲养的窝数。

L = 用周表示的到最后一窝的世代间隔。

$\bar{L}$ = 所有窝的平均世代间隔（见正文）。

p = 选择的比例。

i = 选择强度。

反应的度量

进行了一代或多代选择时，有关实际获得反应的度量引出几个问题。这些是方法而不是原理的问题，将仅作简要地讨论。

世代平均值的变异性

第一个问题由世代平均值的变异性引起。对任何一个选择图解的检查看到世代平均值

不是以一种简单有规则的形式向前进展，而是无规律地且多少有点剧烈地波动着。世代平均值之间这种变异的结果是，在进行若干代选择之前，对选择反应很难做到任何准确度要求的度量。于是每世代平均反应的最好度量，可从与世代平均值吻合的回归直线的斜率中获得，就像下面例子说明的那样，假定真实反应在这一时期内是稳定不变的。

例 11.4

图 11-4 显示了对小鼠体重双向选择 11 个世代的结果（Falconer，1953）。“上选”和“下选”系群分别显示在左图中，两者之间的歧化显示在右图中。对观察到的世代平均值配合了线性回归直线。（选择的第一世代没有考虑，因为选择方法不同。）每世代平均反应的估计值以及它们的标准误如下：

上选	0.27 ± 0.050
下选	0.62 ± 0.046
歧化选择	0.88 ± 0.036

向上和向下选择反应之间的差异将在下一章讨论。

世代平均值变异的原因将在下一章中给予更全面地考虑。这里我们简单地指出原因是什么以及考虑做什么可以减少世代间反应的这种变异。变异的原因是：随机遗传漂变、在估计世代平均值过程中的抽样误差、选择差的差异以及环境因素。由随机漂变和抽样误差引起的变异只有通过增加选择和度量的数目才能降低。选择差的差异可以采用在下一节中说明的方法来减少。世代间环境差异可由很多原因如气候、营养和一般管理等引起。从反应速度的评估中消除环境波动的明显方法是维持一个未经选择的对照群体。根据环境差异对选择群体和对照群体有同样影响的假设，它们之间的差异估计了由选择造成的遗传改良。可是，对照群体的使用并不总是能改善反应估计的精确度。两个群体都受到随机漂变和抽样误差的制约，两个系之间差异的抽样方差是每一个系各自抽样方差的和。而且，实验规模通常受有效设施的限制，这样，对照群体的使用就需要减少选择系的群体大小。如果选择系和对照群体两者都只有一单个选择系群体大小的一半大，那么，按与对照群体的离差度量，对照群体的使用会使反应的抽样方差扩大 4 倍，因此标准误扩大两倍。准确度的这种损失可能抵消了来自消除环境差异的增益。如果对照不是一个未经选择的群体而是一个在相反方向上选择的群体，那么，按所用对照度量的反应的相对准确度能有改善。这被称为双向或分歧选择，在图 11-4 中进行了说明。每一个选择系对另一个系起到“对照”的作用，反应为两个系之间的歧化度所度量。从图 11-4 中清楚看到通过这种方法消除了世代间的一些变异。在世代间缺乏环境差异的情况下，相对于反应大小而言，精确度与全部设备条件相同时用单个系度量的精确度是一样的。其理由是两个系之间差异的标准误加倍了，正如上面所解释的，但选择反应也近似加倍了，因为选择了两个系。可是，由于实践上的原因，如果人们仅对一个方向上的变化感兴趣，则一个未选择对照群体要比双向选择更好，因为两个方向上的反应不总是相等的，这一点将在下一章讨论。

环境的随机变化降低了反应被估计的精确度，但它们不影响估计值。环境的趋向，即随时间的渐进性变化，会引起一个更严重的困难，因为看起来像是选择反应实际上却是由环境趋向引起的。这造成难于评估在改良家畜和植物上选择的效果，因为没有对照，就没

有一种可靠的方法来决定多少改良是由选择引起的，又有多少改良是由改进的管理引起的。可是，像家畜那样，当世代重叠时，根据不同世代的同一年龄个体之间的比较，没有未选择对照群体也能评估遗传改良（见 Smith，1962）。有关反应度量和对照使用的详细考察，见 Hill（1972a）和 Muir（1986）。Hill（1972b）综述了有关对照有用性的实验证据。

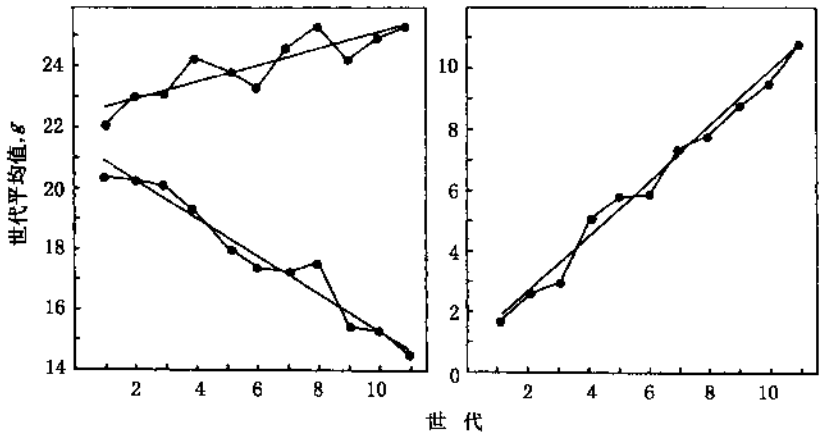


图 11-4　对小鼠 6 周龄体重的双向选择

两个系各自的反应显示在左图。歧化，即上选和下选系之间的差异显示在右图。见例 11.4（根据 Falconer，1953）。

加权选择差

在实验选择中，必须度量选择差和选择反应，因为从遗传学的观点来看，人们感兴趣的不仅仅是反应自己，而正是两者之间的关系。我们必须区分期望的和有效的选择差，因为实践中单个亲本对后代的贡献不等。生殖力的差异总是存在的，因此一些亲本比其他亲本贡献更多的后代。为了获得选择差的度量，它与从后裔世代均值中观察所得的反应有关，因此，我们必须根据被度量的后代数对亲本的离差进行加权。如同本章开头时所下的定义，期望选择差是亲本表型离差的简单平均；有效选择差是亲本的加权平均离差，分配给每一个亲本或每对亲本的权重是亲本对在下一世代中被度量个体的贡献比例。

选择差的权重计入了很大一部分自然选择的效应。如果生殖力差异与亲本被选择性状的表型值有关，那么这种自然选择要么有助于要么有碍于人工选择。例如，如果较极端的表型生殖力较低，或者更经常的不生育，那么自然选择对人工选择起不利作用。通过加权选择差，我们一起度量了自然和人工选择的联合效应。因此有效（即加权）选择差与期望选择差的比较可用来发现自然选择是否在起作用。

例 11.5

图 11-4 显示的小鼠实验在上选方向进行了 30 个世代，在下选方向进行了 24 个世代（见 Falconer，1955）。在表 11-4 中，对两个系中有效（加权）和期望（未加权）的选择差进行了比较。选择的时期分成两个部分，比较在每一部分单独进行。在向上选择的整个时期内，有效和期望选择差之间实际上没有差异，我们能得出结论自然选择作为影响反应的

一个因素是不重要的。可是，在向下选择系中的情形就不同了，有效选择差低于期望选择差，尤其是在第二时期里。从这里我们可以得出结论，自然选择起的作用有利于体重大的家系，因此阻碍人工选择，降低了获得的反应，尤其是在实验的后期中。自然选择对体重小的家系起作用的主要方式是通过生殖力，较小的小鼠其窝中倾向于具有较少的幼鼠，且更经常不育，因此最小的被选择亲本由较少的被度量后代所代表。还有另一种方法可使有效选择差减少，它在形式上等价于自然选择。较小的小鼠比较大的小鼠倾向于更有反应性和易受惊，很小的小鼠在换笼期间经常逃跑和丢失。在它们繁殖之前用这种方式丢失的小鼠，表明是中选小鼠中最小的。

表 11-4　不同选择方向下每世代的选择差

选择方向	世代数	每世代的选择差（g）		
		期望值	有效值	有效值/期望值
上选	1 ~ 22	1.39	1.36	0.98
	23 ~ 30	1.08	1.09	1.01
下选	1 ~ 18	1.03	0.96	0.93
	19 ~ 24	0.82	0.70	0.86

选择差的权重没有计入自然选择的全部效应，因为它没有容纳后裔间存活力的任何差异，而这种差异与它们的表型值有关。

现实遗传力

正如图 11-4 所示，每个世代的反应，描述了什么会发生，但它没有考虑被应用的选择的大小。因此需要一种方法来揭示反应如何与选择差有关。采用下面的方法，通过将反应表达为选择差的一种比例，即比率 R/S 可以做到这一点。总均值以累积的选择差为背景绘制成图。也就是说，经适当加权的选择差对连续世代求和，以便给出对讨论中的那个世代所施加的总选择。用这种方法绘制的反应在图 11-5 中进行了说明。这样，对这些点配合的回归直线斜率给出了比率 R/S 的平均值。

选择反应可用作估计基础群体遗传力的一种方法，因为比率 R/S 的期望值就是遗传力，对等式［11.2］重新整理，有：

$$h^2 = \frac{R}{S} \qquad [11.7]$$

用这种方法估计的遗传力被称为现实遗传力（realized heritability），因为它是反应的一个基本描述，由于下列几个原因，它不提供基础群体遗传力的有效估计值。首先，具有高遗传力性状的反应在第一世代选择之后预计有点下降，这样第一世代之后的现实遗传力将低估基础群体的遗传力，将在本章最后解释其原因。其次，如果存在由环境趋势或近交衰退引起的系统变化，这些将包括在反应中，除非通过与对照群体的比较来消除它们。由随机漂变引起的变化也和反应混杂在一起。下一章将讨论的随机漂变的效应经验上只能通过重复选择来评估。图 11-5 提供了现实遗传力被其他因素干扰的一个明显的例子。在两个方向上的选择产生了很不同的现实遗传力；每一个都是反应的一个有效描述，但它们两个

都不是基础群体遗传力的有效估计值。两个方向上反应不同的原因也将在下一章讨论。基础群体中遗传力的无偏估计值可通过 REML 从选择反应中获得，REML 使用来自选择反应以及每一世代内和世代间所有个体之间遗传关系的信息（例如，Meyer 和 Hill，1991）。REML 是估计现实遗传力的最好方法，因为它考虑了漂变和选择。

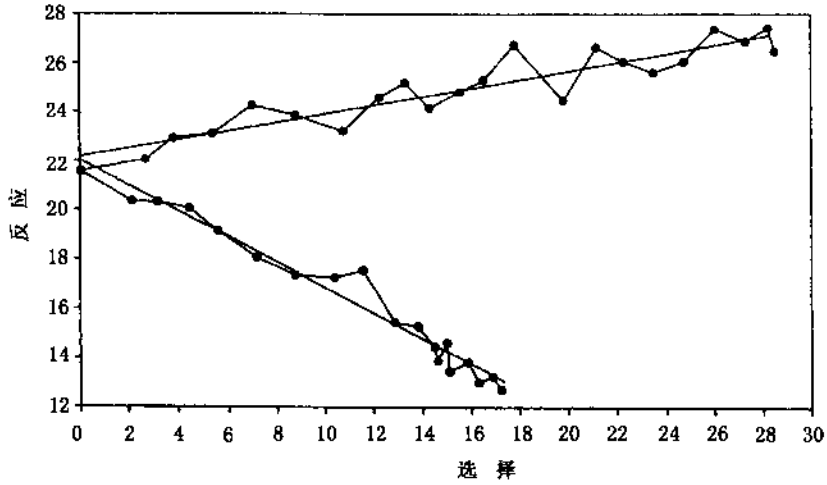


图 11-5 对小鼠 6 周龄体重的双向选择。世代平均值以累积的选择差为背景绘制成图，正如正文中所解释的那样。对这些点配合的回归直线斜率度量了现实遗传力，对向上选择它是 0.175，对向下选择它是 0.518（根据 Falconer，1954）

母体效应

第 9 章描述了作为母亲和后代间相似性一种来源的母体效应。母体效应可导致不规则的选择反应，这些反应比没有母体效应时预测的要大一些或小一些，它们甚至可以导致反应走向“错误”的方向或者适应失调（maladaptive）（Kirkpatrick 和 Lande，1989）。一个性状在“错误”方向起反应的例子是小鼠的窝产仔数（Falconer，1960*a*，1965*b*）。当母鼠被选择具有大窝时，它们的后裔在这些大窝中被哺乳不得不和较大数目的窝同伴分享乳汁。因此，后裔生长更慢，比成年个体要小。较小的鼠倾向产生较小的窝产仔数，因此这种母体效应在与选择相反的方向上起作用。在这种双向实验中，与向下选择的系相比，向上选择的系在第一和第二代有较小的窝仔数。

如果性状受制于母体效应，则在对选择反应做可靠预测之前或来自观察反应中的遗传力估计值能被信赖之前，对母体效应的详细了解是需要的。

人工选择下基因频率的变化

在本章一开始就指出，由选择引起的群体平均值的改变是通过影响被选性状位点上的基因频率变化实现的。但因为位点的效应不能被单个识别，所以实践中无法追踪基因频率的变化。可是，推导将作用于单个位点的选择强度 i 和选择系数 s 关联起来的近似表达式是可能的。于是通过将适当的 s 值代入第 2 章给出的公式中，可以找到基因频率的近似变化。

一个数量性状的选择对相关位点上该性状的影响可以用图 11-6 说明的方式来图示。这指的是具有两个等位基因的一个位点，仅表示了两个纯合基因型 A_1A_1 和 A_2A_2。杂合基因型的位置将在以后考虑。两个基因型具有相等的频率，但这对于论证并非必需。采用先前几章的符号（见图 7-1），纯合基因型其平均表型值相差 $2a$ 个度量单位。描绘了每一个基因型内表型值的分布，这种变异由其他位点引起，也由环境原因引起。假定两个分布都是正态的，并且每一个基因型内方差都是相同的。实心垂直线表示截断点，位于其右边的两个基因型的所有个体都被选择。

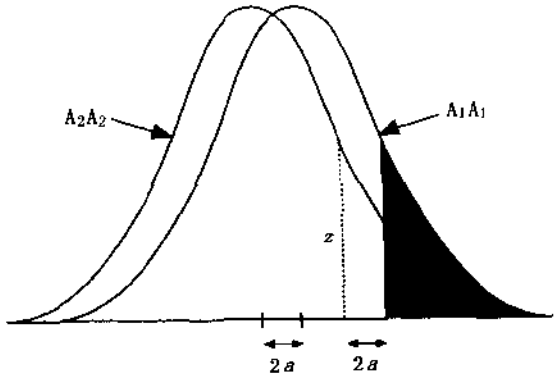


图 11-6　对相关的一个位点上起作用的一个数量性状的选择

通过找出两种基因型的相对适合度，即选择全程中存活的相对比例，可以推导出选择系数。假定 p 是存活的 A_1A_1 的比例，在图中由阴影表示。现在设想截断点移动 $2a$ 个单位，正如虚线所示，这样相同比例 p 被这个新截断点从 A_2A_2 基因型中切掉了。实际上被选择的 A_2A_2 基因型的比例是 p 减去位于两条截断直线之间的部分。只要两种基因型之间的分离即 $2a$ 相对于标准差很小，那么在两条截断直线之间的 A_2A_2 曲线的面积就是 $2az$，其中 z 是纵坐标的高度。与这个面积相对应的 A_2A_2 基因型的比例是 $2az/\sigma_p$，其中 σ_p 是表型标准差。根据等式 [11.5]，$z = ip$，因此实际被选择的 A_2A_2 的比例是 $p - 2aip/\sigma_p = p(1 - 2ai/\sigma_p)$。在第 2 章，选择系数指的是选择不利的基因型下降的适合度，这里它是 A_2A_2。不利于 A_2A_2 的选择系数可从它的相对适合度中获得，结果如下：

$$1 - s = \frac{A_2A_2 \text{ 的适合度}}{A_1A_1 \text{ 的适合度}} = \frac{p(1 - 2ai/\sigma_p)}{p}$$

从中，不利于 A_2A_2 的选择系数可变成：

$$s = i\frac{2a}{\sigma_p} \quad \text{（近似）} \qquad [11.8]$$

虽然在它的推导过程中只考虑了两种基因型，但这种关系对任何程度的显性都是有效的。如果 A_1 完全显性，则杂合子包括在 A_1A_1 基因型内，并且，如果 A_1 完全隐性，则杂合子包括在 A_2A_2 内。在上述任一种情形下，基因频率的改变由等式 [2.8] 近似给出，它是 $\Delta q = -sq^2(1 - q)$，其中 q 是 A_2 的频率。如果无显性，则对杂合子起作用的选择系数在第

2章中被定义为$\frac{1}{2}s$。A_1A_2和A_1A_1之间平均值的差异是a，等价于等式［11.8］的关系是$\frac{1}{2}s = ia/\sigma_p$，它与等式［11.8］是相同的。因此通过将等式［11.8］中的s代入等式［2.7］可以获得基因频率的变化，它是$\Delta q = -\frac{1}{2}sq(1-q)$。Latter（1965）检查了等式［11.8］为s提供一个合理的近似值的条件，更一般性的推导由 Kimura 和 Crow（1978）给出。

等式［11.8］指出了表示选择强度的两种方法（即通过选择强度和选择系数）是如何互相关联的。一个数量性状平均值的变化可以用两种相对应的方法推导，两种推导的等价性为等式［11.8］提供了检验。从s和Δq中可以推导出平均值的变化，结果如下。考虑无显性的一单个位点。根据等式［7.2］，具有$d=0$的平均值是$M = a(p-q) = a(1-2q)$，其中q是起始基因频率。平均值的变化是$\Delta M = a(1-2q_1) - a(1-2q)$，其中$q_1$是新的基因频率。这可以简化成$\Delta M = -2a\Delta q$。将来自等式［2.7］的$\Delta q$代入，给出$\Delta M = asq(1-q)$。等式［11.8］允许我们把$s$变换成$i$的形式，这给出$\Delta M = 2ia^2q(1-q)/\sigma_P$。现在，当$d=0$时，根据等式［8.5］，则有$2a^2q(1-q) = V_A$，其中$V_A$是由所考虑位点引起的加性方差。对所有位点求和给出用总加性方差形式表示的平均值变化为$\Delta M = iV_A/\sigma_P = (iV_A/V_P)\sigma_P = ih^2\sigma_P$，这是由等式［11.3］给出的选择反应。因此，等式［11.8］为从s和Δq中推导出的平均值变化以及从i和h^2中推导的反应之间提供了一种联系。如果作出忽略$(\Delta q)^2$的近似，对一个显性基因，可以证明具有相同的等价性。

等式［11.8］的数量$2a/\sigma_P$是以表型标准差的形式表示的两个纯合子之间数值的差异。这种数量将被称为位点的标准化效应。在下一章中等式［11.8］将被用来对引起选择反应的位点的标准化效应得出一些尝试性的结论。

选择对方差的影响

遗传方差取决于基因频率，如图 8-1 所示。因此可以预料在选择反应中基因频率改变时，遗传方差也发生变化。方差的改变不能预测，但在最初几个世代中其改变很可能很小，因为基因频率变化缓慢，除非存在效应巨大的基因。下一章描述较长时期内的方差改变。

可是，有一种变化原因，它借助最初的选择立即起作用。由于下面解释的原因，选择降低了最初选择的后代方差和遗传力。在随后的世代中减少仍继续但逐渐变小。由选择造成方差的这种下降被称为布尔姆效应（Bulmer effect）。它是如何发生的，简述如下；全面的解释可参见 Bulmer（1985，第 9 章）。

一组经选择的亲本相当于表型分布的一个末尾部分，因此它们的表型方差一定小于从中选择它们的整个群体的方差，正如第 10 章中指出的那样。如果V_P是选择前的表型方差，k是它减少的因子，那么在选择亲本中的表型方差V_P^*为：

$$V_P^* = (1-k)V_P$$

因子k取决于选择强度。当选择通过正态分布的截断进行时，则有：

$$k = i(i-x) \qquad [11.9]$$

其中 i 是选择强度，x 是截断点偏离群体平均值的相应的离差（i 和 x 的值给在附录表 A 中）。选择个体之间的加性方差可以推导如下。育种值 A 和表型值 P 之间的相关是 h（等式 [10.3]）。相关的平方与 A 方差呈比例，而这一比例与 P 的变异相关联。因此，只有加性方差的比例 h^2 受选择个体中下降的表型方差的影响。V_P 减少了因子 k，V_A 的比例 h^2 也减少相同的因子。这样，选择个体之间的加性方差为：

$$V_A^* = (1 - h^2 k) V_A$$

其中，V_A 是选择前群体的加性方差。

加性方差的减少可采用配子期不平衡的形式来描述。这在第 8 章中作为非随机交配的一种形式的结果而提及，并被等式 [8.10] 中的协方差项所体现出来。这里由不平衡引起的协方差是 $-h^2 k V_A$；因为是负的，它降低了方差。用下面的方法，可以理解为什么会产生负的不平衡的理由。假定我们能选择一组个体，它们都具有相同的表型值，并且假定我们能够记录每一个体中每一个基因对性状影响的大小。于是我们会发现基因-效应在个体中是呈负相关的；换句话说，基因—效应之间存在负的协方差（这与在第 9 章结束时有关“竞争”一节中解释的现象基本上相同）。当然，根据其表型值选择的个体不会具有同一的基因型值，但它们的基因型值比随机选择的一组个体更相似，因此由选择产生了负的不平衡协方差。

现在考虑被选择个体随机交配产生全同胞家系时后裔的加性方差，正如第 1 章所解释的那样，由不连锁基因引起的不平衡在后代中减半了，但由于下面的原因，它只出现在家系间方差中。由不连锁基因引起的家系内方差不受亲本表型值的影响，因而也不受选择的影响。这样很简单，它就是 $\frac{1}{2} V_A$。全同胞家系平均值的加性方差是中选亲本的加性方差的一半。因此它是 $\frac{1}{2}(1 - h^2 k) V_A$，不平衡协方差是 $-\frac{1}{2} h^2 k V_A$。将家系间和家系内方差加起来就给出由不连锁基因引起的总加性方差，即 $\left(1 - \frac{1}{2} h^2 k\right) V_A$。上面概述的一个世代选择的结果汇总在表 11-5。

表 11-5 中选亲本及其后裔中的方差组分

组分	选择后亲本	后代(1 世代)	
表型	$(1-k)V_P$		$\left(1 - \frac{1}{2} h^4 k\right) V_P$
加性	$(1 - h^2 k) V_A$	家系间	$\frac{1}{2}(1 - h^2 k) V_A$
		家系内	$\frac{1}{2} V_A$
		总的	$\left(1 - \frac{1}{2} h^2 k\right) V_A$
不平衡	$-h^2 k V_A$		$-\frac{1}{2} h^2 k V_A$

注：V_P 和 V_A 是选择前亲代的方差，h^2 是选择前亲代的遗传力；据等式[11.9]，$k = i(i - x)$。家系是全同胞。

在上面的处理中显性方差被忽略了；但在任何情况下，如果显性方差存在，它将很少受到不平衡的影响，因为它出现在家系间组分中只有 1/4。当然，后代中的环境方差与选

择前亲本中的相同。因为 V_D 和 V_E 未变，或几乎这样，下降的加性方差引起遗传力的下降。如果 h_1^2 和 h^2 分别为后代和亲代的遗传力，两者的比率可根据表 11-5 中的数量来表示，为：

$$\frac{h_1^2}{h^2}=\frac{1-\frac{1}{2}h^2k}{1-\frac{1}{2}h^4k} \qquad [11.10]$$

第一代的选择反应取决于亲代的遗传力，可以用等式［11.3］来预测。第二代的选择反应减少了，因为遗传力和表型方差两者都下降了。如果 R_1 和 R_2 是开始两代的选择反应，根据等式［11.10］和表 11-5，两者的比率可表示为：

$$\frac{R_2}{R_1}=\frac{1-\frac{1}{2}h^2k}{\sqrt{1-\frac{1}{2}h^4k}} \qquad [11.11]$$

这个比率的一些代表值列于表 11-6 中。可以看出，对高遗传力和高选择强度，在第二代中的期望反应值大约比第一代的低 20%。可是，对于低遗传力，这种减少要少得多。

表 11-6 用基础群体的遗传力 h^2 和被选择比例 $p\%$ 的一些代表值，根据等式［11.11］计算的相对于第一代的第二代选择反应

$p\%$	遗传力 h^2				
	0.9	0.7	0.5	0.3	0.1
5	0.76	0.79	0.83	0.89	0.96
20	0.78	0.81	0.85	0.90	0.96
50	0.83	0.85	0.88	0.92	0.97

虽然由于后代中的重组使不平衡（未连锁基因）丢失了一半，但通过下一步选择又会重新产生更多不平衡。任一世代（$t+1$）的加性方差可借助循环等式用前一世代的加性方差表示：

$$V_{A(t+1)}=\frac{1}{2}[1-h_{(t)}^2k]V_{A(t)}+\frac{1}{2}V_A$$

其中 V_A 是基础群体的加性方差（见 Bulmer，1985，等式 9.30）。可是，加性方差的减少不是不确定地发生。平衡很快就达到了，其中由选择引起的不平衡的增加被由重组引起的损失所抵消了。在缺乏连锁的情况下，3 或 4 个世代就足以令群体十分接近于这种平衡，尽管对连锁基因将花费较长的时间。在达到平衡状态之后，预计反应将保持稳定不变，直到基因频率的变化大得足够影响选择反应为止。因此在实验中通常发现的反应的恒定不变与选择对方差的影响并不相矛盾，至少在第一代之后是这样。但是，预计在第一代之后观察到的反应，在某种程度上比从基础群体的遗传力根据等式［11.3］预测的反应要低。当遗传力和选择强度低时，这种差异很小，当它们较高时，这种差异也就更大。

例 11.6

Atkins 和 Thompson（1986）描述过绵羊的一个实验，它为预测因选择造成配子期不平衡的效应提供了一个极好的证据。在苏格兰黑面绵羊群体中对胫骨（后肢一种骨头）的长度进行双向选择，按体重校正了长度。在 19 年内选择大约实施了 8 个世代。从上选和下选系之间的歧化中估计了反应。有一个未选择的对照系，从中估计了基础群体的参数。基础群体的遗传力为 0.56 ± 0.04。将我们不需涉及的复杂性考虑进去之后，根据等式[11.7]计算的现实遗传力是 0.52。正像从配子期不平衡效应中所预测的那样，现实遗传力低于基础群体的遗传力。为了检验观察到的反应是否与考虑不平衡时所预测的反应大小一致，作者计算了计入不平衡时能给出观察反应所需要的基础群体遗传力。他们发现这是 0.57，它与观察值 0.56 吻合得很好，表明由不平衡引起的反应减少几乎与预测的完全相同。

表型方差也减少到了预测的数量。在高系中它减少了 9%，在低系中减少了 11%，而预测的减少量是 10%。

习　题

为简炼起见，本章结束时描述的选择对方差的影响，在所有这些习题中都将之忽略，但 11.3 除外。

11.1　如果选择应用于表中的性状，个体根据它们自己的表型性状进行选择，那么每个世代预期的进展速度将是多大？

	遗传力	表型方差	选择比例(%)
(1)小鼠的体重(g)	0.37	10.7	(*a*)25
			(*b*)50
			(*c*)75
(2)赤拟谷盗(*Tribolium*)的发育时间(天)	0.18	1.7	10
(3)小鼠的雌性生殖力(窝产仔数)	0.22	4.3	30

[解答序号：9]

11.2　习题 10.8 中达尔文雀科鸣禽物种由于 1977 年的干旱而遭受严重死亡。这个物种吃种子，干旱期间可用的种子主要是大而硬的种子。与干旱以前的群体相比较，生存下来的鸟类几个性状的直径，特别是喙的直径较大。干旱前 642 只鸟喙的平均深度是 9.42mm，干旱后 85 只鸟的是 9.96mm。从这个选择性生存和习题 10.8 的数据中，你能预测喙的深度将发生多大变化？

数据来自 Boag，P.T. 和 Grant，P.R.（1981）*Science*，**214**，82～84。

[解答序号：19]

11.3　假定计划对一群肉仔鸡从 5 到 9 周龄的增重进行选择。根据下面的数据，预测选择 5 个世代之后的平均增重。在每一世代中，选择 4 只雄性和 8 只雌性，它们来自被度量的

60 只鸡中。基础群体：平均值 = 738g，标准差 = 111g，遗传力（来自同胞分析）= 0.81。

数据基于 Pym，R.A.E. 和 Nichols，P.J.（1979）*Brit. Poult. Sci.*，**20**，73 ~ 86。

［解答序号：29］

11.4 表中数据指的是提高和降低小鼠血浆胆固醇水平的选择。单独从两个系中和从它们的差异中来计算现实遗传力。*M* = 世代平均值，*P* = 用作下一世代亲本的中选个体的平均值。单位是 log（mg/100ml）。两性进行平均。

世代	高		低	
	M	*P*	*M*	*P*
0	2.16	2.32	2.16	2.02
1	2.26	2.34	2.06	2.00
2	2.26	2.37	2.03	1.97
3	2.33	2.41	2.02	1.96
4	2.45	2.47	2.05	2.01
5	2.44	—	2.01	—

数据来自 Weibust，R.S.（1973）*Genetics*，**73**，303 ~ 312。

［解答序号：39］

11.5 下表给出了图 11-5 显示的实验的最后世代中，对身体大小小的小鼠选择的数据。计算未加权和加权的选择差。单独处理雌性和雄性亲本，然后将它们联合在一起。关于自然选择能得出什么结论？

交配数	亲本的重量（g）		度量的后裔数
	雌性	雄性	
1	7.6	12.4	1
2	12.4	14.5	9
3	13.5	11.6	0
4	13.7	11.6	15
5	13.2	14.3	5
6	17.2	17.1	14
7	10.7	13.8	10
8	12.9	11.6	9
9	14.2	10.1	0
10	10.5	13.1	6
亲本被选择的群体平均值是： 雌性，13.14；雄性，14.80			

［解答序号：49］

11.6 假定在一个实验中你用下面的方法选择绵羊的生长速度。公母等量留种且以一

单个对子进行交配，被选择的数目很大。当它们 2 岁时，公母开始配种，随后每年产一次羔。养育至成熟的平均羔羊数为每个配种季节每头母羊 1.2 头。每年选择相等的数目。为了使每年的进展速度最大，在通过选择后代来更新它们之前，你应当饲养这些亲本多少年？于是，每一年中选择多少比例的羔羊？

[解答序号：59]

11.7　习题 11.6 中的绵羊育种方案实际上是不现实的。为了更实际一些，假定不是每一只公羊只配一只母羊，而是每一只公羊可配 10 只母羊。这影响到最适宜的方法，现在母羊 4 岁时就淘汰，使用了 3 个配种季节，公羊 3 岁时就淘汰，使用了 2 个配种季节。这比单一配对交配的最适宜方法要好多少？为什么两性被淘汰的最适宜年龄比习题 11.6 的要低？

[解答序号：111]

11.8　如果要选择母猪的窝产仔数，选择可以它们第一窝的产仔数或两窝以上的平均产仔数为基础。增加窝数具有增加选择强度和增加遗传力的好处，但它有增加了世代长度的坏处。为了使每年的期望反应最大，最适宜的窝数是多少？采用习题 8.6 中的窝产仔数的重复力 0.409。假定（i）母猪 1 岁时产生第一窝仔猪，随后以 6 个月的间隔产生下一窝，（ii）每窝饲养的平均个体数为 8，（iii）选择的数目是更新亲本所需要的数目。为简明起见，进一步假定（iv）选择的数目很大，（v）世代不重叠，中选个体的后裔全都有它们第一窝仔，其时最年轻的后裔为 1 岁。

[解答序号：121]

11.9　假定影响果蝇（*Drosophila*）腹侧刚毛评分的基因中有一个是加性的，两个纯合子相差 0.3 根刚毛，增效作用等位基因的频率为 0.4。在表型标准差为 2.0 根刚毛的群体中，在两个连续世代对 10% 的最高评分的果蝇选择之后，这个等位基因的频率将是多大？

[解答序号：131]

第 12 章　选择：Ⅱ. 实验结果

在上一章我们看到关于人工选择效应的理论推导只限于群体平均数的变化，严格地说只限于一个世代。通过基因频率的改变，选择改变了群体的遗传特性，而选择的进一步效应却取决于群体的遗传特性。由于单个位点的效应不可知，不能预测基因频率的变化，因此选择反应只有在遗传特性实质上保持不变时才能预测。这样就有许多的选择后果只能通过实验才能发现。本章的目的是简述关于从选择的实验研究中得出的这些后果的最一般结论是什么。实验所要回答的最重要的问题是关于选择的长期反应。如反应可以持续多久？群体均数最终被改变的程度有多大？限制进一步进展的遗传本性是什么？但是，在详述长期选择反应之前，关于早期世代有两个问题需要考虑，在早期世代中反应速度大致保持不变。这两个问题就是在相反方向的选择中反应的重复性和反应的不对称性。

短期结果

反应的重复性

在上一章中我们评论了世代间反应的变异性，并且列出了 4 个原因：由于亲本数目有限造成的随机漂变、估计世代均数中的抽样误差、选择差的变异以及环境因素。现在要考虑的问题是总反应的变异性：如果重复实验，各次结果的一致程度有多大？在将标准误加在现实遗传力上作为基础群体遗传力估计值之前，或者允许作出不同实验之间的比较之前，对这个问题必须做出回答。这里我们只关注这样的选择时期，在此选择期间，反应保持稳定足以对世代均值配合一条直线回归而没有显著误差。当然能够计算所配回归直线斜率的标准误。但这并不告诉我们重复间斜率的变异有多大，因为它没有考虑随机漂变引起的重复间变异。随机漂变引起基因频率改变，这一点反映在世代均值的变化之中。结果，独立选择的重复系以图 3-2 中说明的方式发生漂变分离。漂变引起的变化是累积的，一个世代中的任何变化可传递作为下一世代变化的起点。因为漂变引起变化的累积性质，偏离任何一条直线回归的离差不能包括由漂变引起的所有变异。因此回归的标准误低估了现实遗传力的标准误。这一点示例在图 12-1 中。示例的实验中有 6 次重复。图 12-1（a）显示了作为单个大群体对待时的所有重复的反应。图 12-1（b）显示了 6 次重复中 3 次的反应。为 6 次重复中的每一个都配合了回归直线，它们的斜率给在表 12-1。从图 12-1（a）的单个回归直线或者从表 12-1 的 6 次重复的平均值中可以估计现实遗传力；它们之间几乎没有差异。可是，标准误有很大的不同。单个回归直线的标准误，上选是 0.020，下选为 0.014；从重复间实际变异中计算的标准误，上选是 0.031，下选为 0.046。

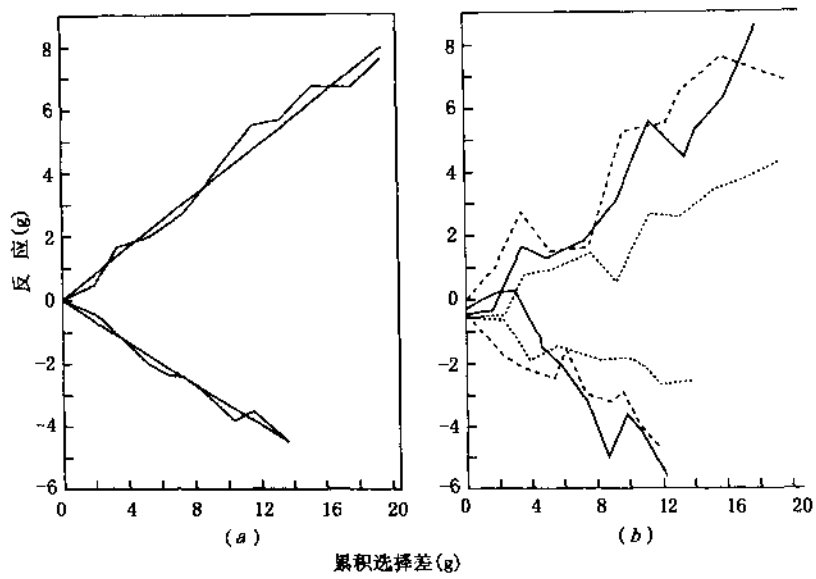


图 12-1　对小鼠 6 周龄体重进行 10 个世代的双向选择（Falconer，1973）。以偏离对照组离差度量的世代平均值对累积选择差（原始数据）绘制成图

（*a*）整个“群体”由每个世代中的 6 次重复选择组成。所配回归直线具有如下斜率 ± 标准误：上选，$b = 0.398 \pm 0.020$；下选，$b = 0.328 \pm 0.014$。

（*b*）每一方向上 6 次重复选择中的 3 次选择。每一重复由具最低近交的 8 对亲本交配产生，给出 $N_e = 31$（由等式［4.9］得来）。这些重复的和剩余重复的现实遗传力列在表 12-1 中。

表 12-1　重复选择系中估计现实遗传力的回归系数*

	上选	下选
	0.457*	0.501*
	0.448	0.376
	0.438	0.365
	0.390	0.301
	0.385*	0.288*
	0.251*	0.159*
平均值	0.395 ± 0.031	0.331 ± 0.046
差异	0.064 ± 0.055	

* 按照它们反应的大小顺序列出重复；显示在图 12-1（*b*）中的那些重复用 * 标记。

抽样方差

获得现实遗传力的抽样方差不是一件直截了当的事。下面的简化说明将允许获得一个近似标准误，详见 Hill（1971，1972*c*，*d*，1980）。重要的结论是回归的标准误严重低估了现实遗传力的标准误；在大范围情况下，后者处于前者的3～5 倍之间。首先考虑在单个

选择系中，用上一代平均值来度量的反应，假定准确无误地知道基础群体均值。假定世代间环境差异可忽略，那么在估计反应的过程中有两种误差来源：由随机漂变引起的方差 σ_d^2 和由度量误差引起的方差 σ_e^2。第一种取决于亲本的有效数目 N_e，第二种取决于估计均值所用的度量个体数 M。没有选择而繁育的一个系的漂变方差大约等于 $2tV_A$，其原因将在第 15 章中解释（见表 15-1）。近交系数 F 大约等于 $t\Delta F$，其中 t 是世代数，ΔF 是近交率，它等于 $\frac{1}{2N_e}$（等式 [4.1]）。因此，$\sigma_d^2 = tV_A/N_e$（近似公式）。选择系的漂变方差并不如此简单，但是相同的表达式可能是一个极好的近似公式（Hill，1980）。来自度量误差的抽样方差简单地就是 V_P/M，其中 M 是被度量的个体数。因此反应的抽样方差为：

$$\sigma_R^2 = \sigma_d^2 + \sigma_e^2$$

$$= V_P\left(\frac{th^2}{N_e} + \frac{1}{M}\right) \quad \text{（近似）}$$

如果反应被估计为具有双向选择或一个选择系和一个对照系的两个均值之间的差异，那么反应的抽样方差是这两个系的方差之和。如果两个系中度量的和繁育而来的个体数相同，则反应的抽样方差是上述表达式的两倍。如果通过将反应除以累积选择差 $\sum S$ 而获得现实遗传力，则其标准误为 $\sigma_R/\sum S$。如果大量的固定数目的亲本作为一单个群体繁殖或分成重复系，对标准误不产生任何差异；可是，有重复系的一个好处是能够获得经验标准误（Hill，1971）。通过将上面表达式包括在反应的方差—协方差矩阵中，经 REML 取得的现实遗传力 h^2 估值解释了世代均值、漂变和度量误差的相关。由两个 h^2 值获得的大约 95% 置信区间给出的对数似然率较最大值小 2（例如，Shaw，1987）。

经过第一世代或前两个世代或者甚至更长时间，选择实验经常不产生反应，但后来产生了明显的反应。原因经常是数目太小，“错误”方向的随机漂变使反应无效。因为漂变的缘故，从经过几个世代选择的小群体中不能获得现实遗传力的可靠估计值。可以近似地计算为获得一给定精确度所需要的实验规模（参见 Nicholas，1980）。

反应的不对称性

图示在图 11-4 和图 11-5 的实验指出了相反方向选择反应的不同速度。当比较现实遗传力时，增加体重的选择只有减少体重选择效率的 1/3。不对称反应已在很多双向选择实验中发现，事实上大部分实验都表现出某种程度上的不对称性。由于下列原因，反应的不对称性有重要的实际后果。从基础群体估计的遗传力中做出反应预测。假定在两个方向上预测反应均值，并且如果存在不对称性，那么一个方向的反应将低于期望值。因此如果一个经济重要性状只在一个方向选择，反应可能会比预期的低而使育种者失望。能够预测不对称性似乎在什么时候，尤其是什么方向上发生，这将是有用的。但是只在一个非常有限的程度上可以做到这一点。原因是引起不对称反应有几个可能的原因，其中只有一些原因能在应用选择之前经观察而被揭示。可以产生不对称反应的主要原因如下。

1. 随机漂变　如果在每个方向上只有一个选择系，反应的不对称性可以容易地由随机漂变产生，像在前一节中解释的一样。因此在任何特定的情形中，第一个问题一定是：在两个方向的现实遗传力截然不同这个意义上，不对称性是否真实。没有选择系的重复就

不容易证明不对称性的真实性。图 11-5 中没有重复，因此其不对称性的真实性是可疑的。图 12-1（a）中有重复并且证明不对称性不多于偶然预测的。由随机漂变产生的不对称性是不能预测的。

2. *选择差*　在上选和下选系中选择差可能不同，这有几个原因。(i) 自然选择可能在一个方向上辅助了或在另一个方向上阻碍了人工选择。(ii) 生殖力可能改变，因而导致在一个方向上比另一个方向上获得较高的选择强度。(iii) 作为均值变化的结果，方差可能改变：选择差随着方差的增加而增加，随着方差的减少而减少。这是一种“尺度效应”，将在第 17 章更全面讨论。选择差的不同可影响每个世代的反应，以及观察和预测反应间的一致性，但它们很少影响现实遗传力（Falconer，1954）。因此，现实遗传力的不对称性不能归因于任何通过选择差起作用的原因。

3. *近交衰退*　有关选择的大多数实验是用偏小的群体进行的，因此在选择进行过程中经常有可测数量的近交产生。如果被选择性状是受制于近交衰退的一个性状，其平均数将有一种因近交而减少的趋势。这将会减少上选方向的反应率和增加下选方向的反应率，因此引起了不对称性。受制于相同近交衰退的未选择的对照群体将揭示有多少不对称性可归因于此原因。对近交衰退率的预先了解将有利于预测反应的不对称性。

4. *母体效应*　被母体效应复杂化的性状可以指明与该性状的母体效应组分有联系的反应不对称性。图 11-5 所示实验中的不对称性就是这种情况。被选择的性状（小鼠 6 周龄体重）可被分成两个组分，断奶重和断奶后生长，前者是由母体决定的。断奶重在大家系中几乎不增加，而在小家系中却下降许多。这样，反应的不对称性全在性状的母体组分中。可是，把反应的不对称性归因于母体效应，只是把问题从选择性状转移到另一个性状上，并没有解释不对称性。

5. *遗传不对称性*　加性遗传方差和遗传力取决于基因频率。当基因频率等于 0.5 时，加性基因对遗传力贡献最大，而当隐性等位基因频率等于 0.75 时，隐性基因对遗传力贡献最大（见图 8-1）。这些将被称为“对称”基因频率。如果影响性状的所有基因在起始群体中都有对称频率，那么随着基因频率的变化，现实遗传力将逐渐减小，而且在相反方向选择的系群中，这种减小量大致相等，因而没有不对称性。可是，假定群体以高于或低于这些值的基因频率开始，那么在一个系群中基因频率将偏离对称值，遗传力将减小。但在相反方向选择的系群中，基因频率将向对称值移动，遗传力将增加。因此，当基因频率在上选和下选系群中变得不同时，不对称性才能出现。观察到的反应取决于所有位点的联合效应，如果“平均”基因频率与加性基因的对称值 0.5 和隐性基因的对称值 0.75 不同，预期将出现不对称性，所谓“平均”是基因效应的加权。然而，不能期望在开始的几代中立即出现由这种原因引起的反应不对称性，因为它取决于基因频率的分化。此外，它与非线性反应有关，因为它取决于一个系群中的反应减速和另一系群中的反应加速。因此它难以解释不能检测的非线性反应的不对称性。可按起始点与选择极限的关系不同方式审视遗传不对称性。所有的有利等位基因被固定时，就是选择的理论极限。如果起始群体表型值不在两个极限表型值中间，反应的不对称性将会出现，以致于选择反应在一个方向比在另一个方向上前进得更远。如果选择有利于杂合子，则情形有点不同，因为在一个方向上的极限不是固定的而是平衡的基因频率。如果起始群体涉及的基因频率不在加性方差最大的这点上，就会产生不对称性。

6. 巨效基因 在第一代中立即出现的反应不对称性能产生于巨效基因的遗传不对称性。不对称性立即产生的原因是亲本的首次选择产生了巨大的基因频率变化，等价于对微效基因的多世代选择。不对称性是由不在对称点上的起始基因频率引起。如果开始反应就是不对称的，那么它表明基础群体中后代对中亲值的回归将是非线性的（Robertson，1977*b*）。因此应当能够预测这种不对称反应。

7. 标量不对称 遗传和环境的变异可能偏斜到不同程度或处在相反方向中。于是在分布的一端比在另一端遗传变异占总变异的比例较大。结果，基础群体中亲子回归将是非线性的，且在第一代中反应是不对称的。面对的情形形象地展示在图 12-2（*a*）中，其中遗传和环境方差在相反方向上偏斜。偏斜度不同可能是尺度效应，这一点将在第 17 章中解释，或者它可能按下列方式由基因型与环境互作引起。经历好环境的个体比经历较差环境的个体可以表现出较少的遗传变异；这图示在图 12-2（*b*）中。或者，具高遗传值的个体比低遗传值的个体可能对环境的变异更敏感，正如图 12-2（*c*）所示。两种情形下，高值个体比低值个体具有较低的遗传力。偏斜度的差异与图 12-2 显示的另一种方式相吻合，在那种情形下，上选的遗传力比下选的遗传力大。而且，这种形式的不对称反应可从基础群体内非线性的亲子回归中来预测。详细内容见 Robertson（1977*b*）。

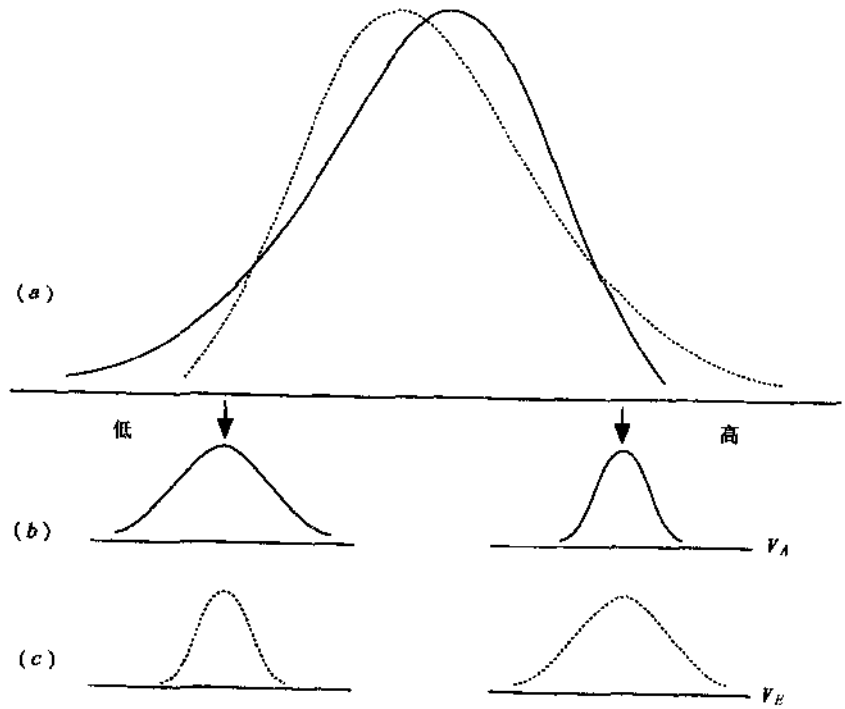


图 12-2 说明标量不对称性的频率分布

（*a*）展示出加性遗传变异（实线）具有负的偏斜度，环境变异（虚线）具有正的偏斜度。（*b*）展示了环境值处于箭头标明的那一点上的个体之间的加性遗传变异。（*c*）展示了有种值处于箭头标明的那一点上的个体之间的环境变异。两种情况下上选比下选将给出较低的现实遗传力。这些图是粗略地按比例绘制的，如果（*b*）和（*c*）中的分布表示选择的个体，那么上选遗传力大约为 0.1，下选遗传力大约为 0.4。

8．*间接选择*　有时选择标准与为估计反应而度量的性状很不相同。于是如果被度量性状与选择标准没有线性关联时，反应的不对称性就可能产生。Baptist 和 Robertson（1976）通过引诱果蝇（*Drosophila*）尽可能远地爬过一系列直径不断缩小的缝隙来选择其体型大小。选择标准是穿过的缝隙数量，这将对称地起反应。可是，这种方法不仅选择了体型大小，而且也选择了活动性。小果蝇比大的更灵活，这导致体型大小和缝隙范围之间一种非线性的联系。结果，体型大小对于上选比下选起反应较少。

由于所有这些可能的原因，经常发现不对称反应是不奇怪的。很难识别在一特定情形中起作用的原因也是不奇怪的。一些原因使不对称性的预测成为可能而另一些原因则令其不能，只有更好地理解不对称反应，从基础群体的遗传力来预测反应率才有一定的可靠性。然而，可以作出一种归纳。它是这样，如果选择的性状是自然适合度的一个组分，由于增加适合度的选择比降低适合度的选择产生较慢的反应，那么应当预测出不对称性。这种期望有 3 个原因。首先，这些性状通常表现出近交衰退，尽管未选择的对照组将揭示这种原因。其次，过去经受自然选择的性状很可能具有频率大于其对称点的有利等位基因。第三，近交衰退表明定向显性导致了非线性亲子回归，尤其是如果存在巨效基因时。对增加适合度而起反应较少的不对称反应的预计已为实验所证实。对果蝇、赤拟谷盗、小鼠、鸡和鹌鹑已发表实验的一项调查（Frankham，1990）发现 30 个实验中有 24 个实验证实了这种期望，不对称性处于相反方向的 6 个实验很容易用实验的特殊特征解释。近交不是有适宜对照的实验中出现不对称性的原因。

长期结果

至少按我们现在的理解能力，长期选择的结果是不可预测的。这有两个原因：首先，结果取决于对反应有贡献的单个基因的特性，这种特性在开始时不能通过观察来确定；其次，因为突变产生了我们无法预测其性质的新变异。然而，有一种理论允许我们预测如果某种条件存在将会发生什么，进而得出什么是实际发生的可能原因。现在我们简述这一理论。我们首先处理在没有突变情况下的反应，然后我们考虑突变的效应。实验中共同发现的一些特征显示在图 12-3 中。当它们对理论有影响时，我们将评论这些特性。

选择极限

没有突变产生的新变异，不可能期望选择反应无限地持续下去。通过选择或伴随着的近交，基础群体中分离的基因迟早会被固定（或如果有超显性会达到平衡）。因此反应会慢慢减少并最终停止。于是群体处于“高原”或*选择极限*。这种类型的长期结果实例展示在图 12-3（*b*）和（*c*）。已发现几个选择实验中逐渐减少的反应与指数曲线吻合得很好（Bünger 和 Herrendörfer，1994；Bünger，Renne 和 Dierl，1994）。这样一条拟合曲线的参数为起始的现实遗传力和处于选择极限的群体均值提供了较好的估计值。关于选择极限提出的问题是：与起始变异有关的总反应为多大，达到极限要多久？首先，一些经典的答案来自于实验。由于不对称性的复杂性，我们只考虑双向选择。当两个歧化系都处在其极限时，它们之间的差异就是总反应 R_T；称之为*范围*。由于随着极限的接近反应逐渐减少，达到极限所需的世代数不易确定。可是，时间尺度能较准确地表达为*半存留期*（half-life），这是达到极限一半时所需的世代数。4 个实验结果汇总在表 12-2 中；两个涉及果蝇，两个

涉及小鼠。达到极限时两个均值的比率变化很大。小鼠高系的窝产仔数只是低系的 1.6 倍（第 4 栏），但果蝇高系的刚毛数是低系的 8 倍（第 1 栏）。但是范围变化从 3.6 到 28 个加性遗传标准差，或者从 1.7 到 20 个表型标准差。所有实验大致经历了 20～30 个世代达到极限，半存留期从 7 个世代变化到 12 个世代。

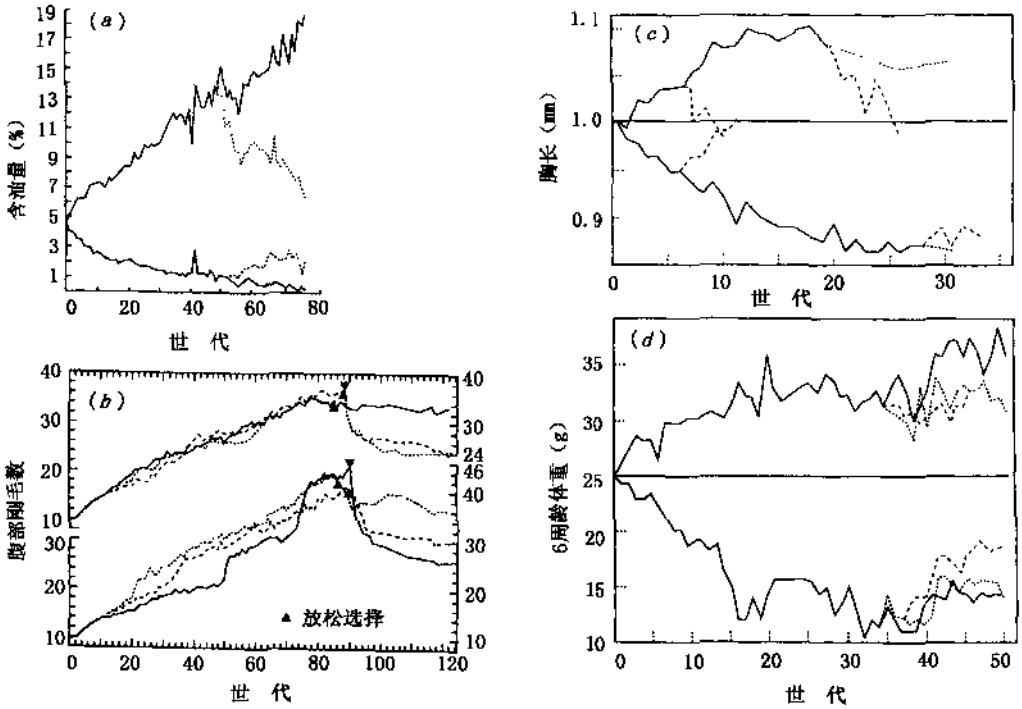


图 12-3　说明长期反应的 4 个实验

（*a*）对玉米种子含油量的双向选择。虚线是相反选择（根据 Dudley，1977）。

（*b*）向上选择腹部刚毛数的黑腹果蝇（*Drosophila melanogaster*）的 6 个重复系。在标记处选择暂停（根据 Yoo，1980*a*）。

（*c*）黑腹果蝇（*Drosophila melanogaster*），胸长（根据 F.W.Robertson，1955）。

（*d*）小鼠，6 周龄体重（引自 Roberts，1966*b*）。

短条线是相反方向的选择反应；点线是人工选择暂停时对自然选择的反应。

（经作者和出版商许可，上述来源的所有图都被重新绘制。）

在这些以及大多数其他实验中，选择已使平均值远离基础群体变异的范围。根据起始群体中现存的变异来计算，虽然表 12-2 中的总反应可能给人印象深刻，但与家畜育种者的成就相比这些总反应一点也不吸引人。例如，在选择达到两个极限之后，小鼠中大的体重是小的 2.5 倍。与之相比，最大品种狗的体重大约是最小品种的 100 倍（Stockard，1941）。当评论家畜品种间差异时，这个令人失望的实验选择结果的原因是实验仅在闭锁群体内进行的，通常这些群体不是很大。相比之下，家畜育种者在寻找有利于他的目的基因过程中，通过断续杂交可以把他的搜寻网撒得远得多，并能利用在别处已经产生的一些可能具有巨大效应的新突变。

那么，关于长期选择，理论能告诉我们什么呢？

表 12-2　4 个实验中选择的极限

		实　验			
		(1)	(2)	(3)	(4)
均值的观察比率	H/L	8.0	1.3	2.5	1.6
观察范围	R_T/σ_A	28	20	16	3.6
	R_T/σ_P	20	12	8	1.7
理论最大值	$R_{(max)}/\sigma_P$	82	52	22	8
有效群体大小	N_e	28	28	15	32
持续期(世代)：	总值(近似)	30	20	25	20
	半存留期	10	7	9	12
	半存留期/N_e	0.4	0.3	0.6	0.4
观察值/最大值：	范围	0.24	0.22	0.36	0.21
	半存留期	0.29	0.19	0.43	0.36
$n = R_T^2/(8\ \sigma_A^2)$		98	50	32	2
标准化效应：	$2a/\sigma_P$	0.21	0.23	0.24	0.95

实验和来源

(1) 果蝇（*Drosophila*），腹部刚毛：Clayton 和 Robertson (1957)。

(2) 果蝇（*Drosophila*），胸长（图 12-3*c*）：F.W.Robertson (1955)。

(3) 小鼠，6 周龄体重（图 11-5）：Falconer (1955)，Roberts (1966*a*)。

(4) 小鼠，窝产仔数：Falconer (1965*b*，1971，1977)。

假定果蝇（*Drosophila*）实验中群体大小为 $N_e = 0.7N$。

极限理论　总反应与起始遗传变异的比例 R_T/σ_A 主要取决于对变异有贡献的位点数。例如，如果只有一个基因频率为 0.5 的加性位点，那么在基础群体中每个最极端基因型出现的频率是 1/4；或者如果有两个这样位点，频率则为 1/16。于是，由最极端基因型代表的选择极限将会完全处于基础群体中发现的变异范围内。对于较多的位点，最极端基因型在基础群体中很稀少，而且选择极限会以 σ_A 为单位从起始均值向远处移动。很清楚，表 12-2 中至少 3 个实验中有很多位点而不是“几个”位点对基础群体的变异有贡献。因为基因数目与基因效应具有反向联系，对于给定数量的遗传变异，如果基因数目很少，则基因一定是巨效的，如果基因数目很大，则基因就是微效的。因为基因数目和它们效应预先都不知道，所以预测极限是不可能的。然而预测极限的理论最大值是可能的，因此让我们在极限理论研究上再深入一点，首先查明极限如何取决于位点数，然后看看极限理论最大值能告诉我们什么。

首先，假设进行选择的群体由相当大量的亲本繁殖而来，因而不会发生随机固定；再假设没有超显性位点。于是处于极限时，所有相关位点上的有利等位基因将成为纯合的。在第 7 章的符号中，范围是 $\sum 2a$ 度量单位，即全部位点上纯合子差异之和。于是范围与起始加性遗传方差如何发生联系呢？为了用一个简单的方法来表示这种关系，我们必须作出两个假设，尽管它们无疑是不正确的，但以后我们会明白误差是如何影响结论的。第一个假设是所有位点对选择的性状具有同样大小的效应。于是范围就是 $R_T = 2na$，其中 n

是位点数，每个位点都有 $2a$ 的纯合子差异。第二个假设是所有基因开始的频率都为 0.5。有了这些假设，通过等式［8.7］，起始加性方差就是 $\sigma_A^2 = \frac{1}{2}\sum a^2 = \frac{1}{2}na^2$（这里使用符号 σ_A^2 而不用 V_A，因为当包含标准差 σ_A 时，它简化了计算式）。注意到基因频率为 0.5 时，显性离差 d 不出现在 σ_A^2 的计算式中，因此不必要作出关于显性度的假设。将范围平方就可获得范围和加性方差之间的关系，这给出：

$$\frac{R_T^2}{\sigma_A^2} = \frac{4n^2a^2}{\frac{1}{2}na^2} = 8n \qquad [12.1]$$

在以后这个等式将被视为估计位点数的一种可能方法。Hill 和 Rasbash（1986）检验了关于基因效应和基因频率分布的其他假设的期望反应。

上面考虑的总反应取决于很大的有效群体大小。实践中使用的亲本数很少多到足以忽略随机漂变的程度。因此发生了一些近交，这导致在一些位点的随机固定。换句话说，不利等位基因在一些位点上固定了，尽管选择对它们是不利的，由此产生的结果是，总反应低于等式［12.1］所指明的。于是所达到的极限取决于每一个位点上有利等位基因被固定的机会，这个机会部分由选择以及部分由近交来决定。近交和选择互相作用的方式已被 Robertson（1960）、Hill 和 Robertson（1966）以及 Robertson（1970）提出。主要结论如下。

位点数通过对每一位点起作用的选择系数 s 来影响这个问题。位点数越多，它们的效应就越小，选择系数也越小。一个有利等位基因固定的机会取决于它的起始基因频率；它越稀少似乎就越易丢失。给定起始基因频率，固定的机会就是 $N_e s$ 的一个函数。$N_e s$ 是有效群体大小和有利于等位基因的选择系数的乘积。选择系数等于 $i(2a/\sigma_P)$，其中 i 是选择强度（等式［11.8］）。因此对于给定的起始频率和给定的基因效应（$2a/\sigma_P$），一个有利等位基因固定的机会是有效群体大小和选择强度的乘积 $N_e i$ 的函数。因此总反应应当随群体大小的增大而增大，随选择强度的加强而增大。这种预计已被一个关于果蝇的实验所证实（Jones，Frankham 和 Barker，1968）。在具有不同选择强度和不同群体大小的许多系群中，选择进行了 50 个世代以增加腹部刚毛数。产生最大总反应的系群是具有最大群体大小和最高选择强度的那些系群。在实践中，提高选择强度经常有必要减少作为群体大小的亲本数目。最好的折衷方法是选择 50%。虽然这样做进展速度会比更强烈的选择慢，但这使得总反应最大（Robertson，1960；Jódar 和 López－Fanjul，1977）。

当群体大小不大，继而存在某种程度的近交时，位点数越多，与起始方差有关的效应越大仍是正确的。如果位点数非常多，那么它们的效应一定很小，在选择能固定有利等位基因之前，大多数位点通过随机漂变被固定下来。尽管这样，如果遗传方差是由相当大量的位点（严格地讲是无穷数目）引起的，即使它们中只有一小部分被选择固定下来，仍可获得最大反应。这就是理论最大反应；如果遗传方差是由无穷数目的位点产生的，则将获得的就是总反应。对于加性基因，理论最大反应可表示为（Robertson，1960）：

$$R_{(\max)} = 2N_e i h^2 \sigma_P \qquad [12.2]$$

这个表达式中所有的项都能估计，因此最大反应能够预测。注意 $ih^2\sigma_P$ 是一个世代的预测反应（等式［11.3］），或是开始几个世代的观察反应。简单地把 i 作为两个方向选择强度

的和就能得到歧化选择的最大反应。然而，对于隐性基因，最大反应可能要大得多，尤其是如果隐性有利基因的起始频率低。如果所有的基因都是加性的，理论最大反应具有 $1.4N_e$ 世代的半存留期；如果基因是隐性的，将会达到大约 $2N_e$ 世代。

理论上的最大反应并不是最佳选择方法所能达到的反应。所达到的反应取决于基因的数目。理论的最大值可在最有利的遗传条件下达到，这是个巨大的基因数，实际上是不可能达到的大数目。因此作为预测，理论的最大值并没有我们期望的上限那么高。然而，将观察反应与它们的理论最大值进行比较有一些好处，因为比较可以对作为理论最大值基础的假设在多大程度上是有效的提出某些看法。如果观察反应比最大值小得多，则观察反应不可能由具有相同效应和相同基因频率的相当多位点产生。关于范围和半存留期的观察反应与最大反应的比率给在表 12-2。比率大约在 0.2 和 0.4 之间，表明这些反应中没有一个是由具相同效应的相当多的基因产生。

近交而不是选择引起的固定程度对处于极限的重复选择系之间的差异有影响。如果近交产生许多固定，则不同的重复将有不同的等位基因被固定在很多位点上。于是重复将达到不同的极限，而且如果它们杂交，一些遗传方差将会被恢复，依此选择进一步起作用。如果近交没有产生很多固定，选择将会在大多数位点上固定相同的等位基因。于是所有的重复将会有相同的极限，并且它们之间杂交将不会产生遗传变异。

表型方差　由上面理论简述期望的遗传方差的损失，应当引起表型方差的减少。可是，很少发现表型方差像期望的那样减少；却是经常增加的。表 12-2 中关于果蝇腹部刚毛数的实验提供了一个例子。图 12-4 显示的频率分布说明了基础群体中以及高重复系和低重复系最极端处的表型方差。两个方向上的选择导致方差大幅度增加。很多其他实验也发现类似的增加。下面是表型方差增加或不减少的一些可能的原因。首先，许多性状的方差不是与平均值无关联的：当在选择下平均数变化时，方差自动地随之改变。这是一种“尺度效应”，将在第 17 章详细论述。其次，环境方差可能增加。随着固定的接近，纯合子频率将增加。有证据表明纯合子由环境原因引起的变异，有时候比杂合子的要大，这在第 8 章已提及，还要在第 15 章全面地讨论。第三，如果存在基因型与环境互作，预计选择将增加环境方差，其理由将在第 19 章中解释。由这些原因引起的表型方差的增加可能平衡了遗传力的下降，因此保持大致稳定的反应率，这一点可从等式［11.3］中看出。第四，遗传方差没有像期望的那样下降，其理由本章后面部分概述。

突　变

Hill（1982*a*，*b*）以及 Hill 和 Keightley（1988）提出了突变如何影响选择反应的理论；这里只给出结论概要。突变不断产生新的遗传变异，但每一个新的突变体在其发生之后只对下几代产生微小的影响。这是因为新的突变基因频率非常低，等于 $\frac{1}{2N}$（其中 N 为群体大小）。在这么低的频率下，新的突变基因对遗传方差贡献甚微，而且如果没有选择，大多数新的突变体将由于随机漂变而消失。可是，选择能增加有利突变体的频率，以致于它们不会消失。因为它们的频率随连续选择而进一步增加，所以它们产生的方差将增加并且对反应有较大的贡献。新的突变在每一代都可引入，这样由突变引起的每世代反应将随着时间推移逐渐积累最终达到一个稳定值，这个值是不可忽略的。当达到这个稳定值时，由

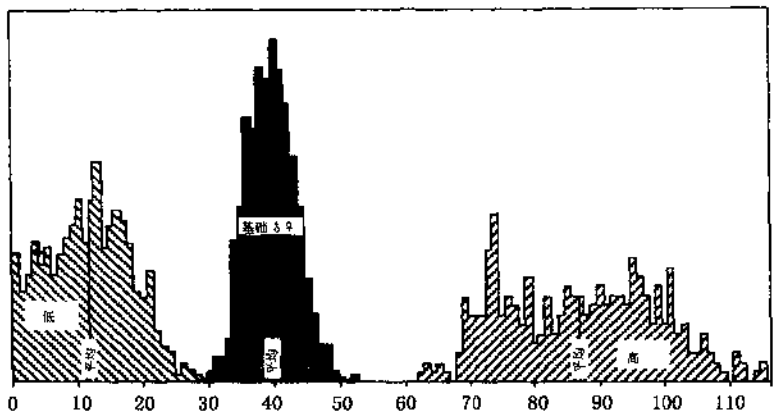


图 12-4 雌性黑腹果蝇（*Drosophila melanogaster*）在基础群体中以及选择 35 和 34 世代之后在最极端高和最极端低的系群中（雌性）腹部刚毛数的频率分布（根据 Clayton 和 Robertson，1957）

连续产生的新突变引起的方差增加就由选择或近交固定的突变基因引起的方差损失所平衡了。在突变开始对反应有可度量的贡献之前，可能要经过 20 个世代，而且在反应率保持不变之前，所花费的时间要长得多。这样突变仅对长期反应是重要的。这有个实际启示：试图用人工诱导突变来改善反应率预计是不会取得任何成功的，除非诱变和选择都持续相当多的世代（Hill，1982*b*）。

当反应率保持不变时，它取决于有效群体大小 N_e、选择强度 i、表型标准差 σ_P 和每一世代由新突变体产生的加性方差 V_M。如果对突变体效应变化的方式（基因效应的分布）作出一些合理的假设，并且如果基因是加性的（无显性）以及基因不影响自然适合度，那么预测的每世代的反应为：

$$R = 2N_e iV_M/\sigma_P$$

（这类似于等式［11.3］，可写成 $R = iV_A/\sigma_P$。）这样在大群体中反应率比小群体中的要大，但达到最大反应率需要较长时间，因为在大群体中，每一个突变基因的频率开始时都较低。突变方差 V_M 只能由实验来确定。对果蝇（*Drosophila*）的一些实验表明，腹部和腹侧刚毛数的突变方差大约为 $V_E \times 10^{-3}$（见 Hill，1982*b*；Lynch，1988）；V_E 是环境方差，它提供了一个便利的标准。这使得测验果蝇长期选择中期望反应与观测反应成为可能。图 12-3（*b*）所说明的实验中，由突变产生的预测反应率是每世代 0.3 根刚毛，而在50～80 个世代间的观察反应率是每世代 0.4 根刚毛（Hill，1982*b*）。这样，这个实验充分支持了突变如何对长期反应有贡献的理论。由突变引起的新变异的产生将在第 20 章中予以更详细讨论。

在长期选择中计入突变时，长期选择发生的一般情况如下。开始，反应来自于基础群体中现有的加性遗传方差。随着加性方差的耗尽，反应率应当逐渐减少。于是大约经过 20 个世代，由突变基因产生的新方差开始对反应有贡献，反应减少速率变小。在经历又一段时间之后，基础群体中分离的所有基因应当因选择或伴随着的近交所固定，于是进一步的反应完全取决于在选择期间已经积累起来的突变。然后反应应当不确定地以一个固定

的速度持续下去：当存在影响被选性状而不是适合度的突变时，则预期没有选择极限。一些实验已经表明反应可以持续相当多世代，而没有接近极限的任何迹象。这些实验中的一个图示在图 12-3（*a*）中，对玉米含油量的选择实验持续到所显示的点之外，直到 90 个世代高油系还没有显示出达到极限的迹象（Dudley 和 Lambert，1992）。另一例是选择赤拟谷盗（*Tribolium*）高蛹重，至少 75 个世代继续有反应（Enfield，1977）。并且图 12-3（*b*）中 6 个重复系中的一个在 88 世代仍有反应。然而，大多数实验已经看到反应终止于一个极限。而且，虽然反应已经停止，但处于选择极限的大多数群体并不缺乏遗传方差。因此我们必须进一步考虑这些事实的可能原因。

选择极限的原因

我们必须区分两种情形：没有遗传方差继续存在时的极限和有遗传方差存在但群体没有反应时的极限。在第一种情形中问题是为什么突变没有引起反应。理由简单地就是由突变产生的新变异不足以产生可察觉的反应；或者可能是大多数突变体对自然适合度有不利影响。存在遗传方差的第二种情形是一种普遍的情形，那么为什么群体没有反应?

遗传方差的存在可以通过颠倒选择方向、或者暂停选择，以便让自然选择单独起作用，或者通过近交来查明。前两种测验应用于图 12-3（*c*）和（*d*）中的选择系。图 12-3（*c*）中的低系对两个测验都不产生反应，表明没有遗传方差存留。图 12-3（*d*）中的高系情况与之相同，但这里向上选择系后来产生了补充的反应，它导致一个新的较高的极限。这可能由新的突变所致。然而，图 12-3（*c*）中的高系和图 12-3（*d*）中的低系对反向选择有反应，表明加性遗传方差是存在的。

下面是遗传方差存在却没有反应的一些可能的理由。

1. 极限可能是由性状的属性或它被度量的方式所强加的一种外在表现。例如：百分数有 0 和 100 两个极限；果蝇的刚毛数不能低于 0。在有关基因几乎没有被固定的情况下，这些极限可能达到或紧密地接近。图 12-3（*a*）的实验非常清楚地显示了这种类型的极限。低系的含油量接近零点的极限，而高系却没有这种外在的极限，继续有反应。

2. 通常在选择系中生殖力减小，并且选择差可能在后来世代中剧烈减小。于是，平均值对世代的图解显示了明显的接近极限，但平均值对累积选择差的图解显示极少有证据表明系群实际上已经停止反应。图 11-5 说明了这一点，特别是在低系中更明显。在这种情况下，系群对反向选择有反应，尽管需要一段时间才能做到这一点。

3. 有利等位基因可以是显性的。当隐性不利基因具有低频率时，它们引起的大部分方差是非加性的（见图 8-1）。在这种情况下，平均数将向对近交不利的方向变化，理由将在第 14 章中解释。小鼠窝产仔数的两个实验的情况被认为存在稀有隐性不利基因，一个见表 12-2（Falconer，1971），另一个由 Eklund 和 Bradford（1977）描述过。在这两种情形中，人们认为近交和杂交已经消除了不需要的隐性基因，因此产生了显著高于原先极限的平均数。

4. 被选择性状的一些位点上存在超显性。处于选择极限时，超显性基因多以中间频率处于平衡之中。它们引起的方差只是非加性的。平均数将有一种趋向近交基础水平的即刻变化。然而，没有证据表明这是选择极限的共同特征。

5. 人工选择可能和自然选择相反。这应当从有效（加权）选择差低于期望（未

加权）选择差中检测出来。选择差的这种损失在例 11.5 已有说明。自然选择也可通过暂停人工选择而被检测出来。图 12-3（*c*）和（*d*）中所有系群的选择都被暂停或“放松”。图 12-3（*c*）中高系可能例外，所有的系都没有反应，表明自然选择不是极限存在的一个原因。

6. 自然选择通过人工和自然选择的共同作用可能有利于杂合子。与被选择性状的简单超显性相反，这种情形通常发生。这种情况可来自适合度上具有有害效应且影响被选择性状的基因，该基因存在于起始的基础群体中，其频率因选择或选择过程中产生的新突变而增加。上面概述的新突变对选择反应贡献的理论假定，影响被选性状的突变就其适合度而言，是中性的。可是，大多数突变很可能是有害的。于是这些突变将不被人工选择所固定，并且选择极限将由人工选择和自然选择之间的平衡所产生，在这种选择极限上遗传方差将继续存在。第 7 章提及的小鼠侏儒基因就是个例子。该基因来自一个小体型中选系的突变（MacArthur，1949），其杂合子被选中，因为它们体型较小。纯合子尽管更小，但自然选择阻止了基因的固定，因为纯合子是不育的。这样，在人工和自然选择的共同作用下，杂合子有优势。在这种情况下，达到选择极限时，就会有基因产生的遗传方差但没有进一步的反应。如果放松人工选择，系群对自然选择就有反应。如果选择被颠倒，人工和自然选择在同一个方向起作用，则群体立刻有反应，通常是快速反应。这可以作为不对称反应的一种极端形式来看待。

经常发现在被选择的果蝇系群中自然选择通过致死性而不是不育性起作用。图 12-4 中说明的选择高刚毛数的系群中，存在这样一个基因，它在纯合子中是致死的，在杂合子中它能增加 22 根刚毛数，这是起始表型标准差的 5.8 倍。发现图 12-3（*b*）中 6 个系群有 5 个含有对刚毛数具有巨大效应的致死因子，其效应可达 $5\sigma_P$（Yoo，1980*b*）。这些致死性解释了选择暂停期间平均值的下降。这些致死性由选择期间的突变所产生，它们解释了在一个系群中看得特别清楚的加速反应。Yoo（1980*c*）进一步分析了实验结束时这些系群体的遗传变异属性，发现从亲属间相似性中估计得到的遗传力，在一些系群中与基础群体中一样高。

与实际育种的关系 可能认为长期连续选择的实验研究与动植物改良的实践很少有关联，因为育种者只关注短期反应。可是，家畜品种已经选择了相当长的时间，在肉仔鸡中，最近选择已经很强烈了。如果它们的表现像一些实验群体，它们现在可能已达到选择极限并不再对选择有反应了。如果真是这样，对选择极限特性的理解将与克服选择极限以及取得进一步进展的方法探究极有关联。可是，幸运的是，主要家畜物种还似乎没有达到选择极限，因为对它们还在做连续的改良（Smith，1984，1988；还可见下一章的最后）。关于实验室实验对家畜育种影响的讨论，见 Roberts（1967*a*，*b*）、Falconer（1971）、Eklund 和 Bradford（1977）、Hill（1982*b*）以及 Eisen（1980）对小鼠实验的综述。

位点（有效因子）数目和标准化效应

因为总反应主要取决于位点数，所以这引导人们试图使用观察反应来估计对总反应有贡献的位点数。可是，解释这样获得的任何数字以及在这一段落中集中解释每个

位点的含义，存在严重的困难。主要有两个困难，首先，影响数量性状的基因没有相等的效应；第 21 章描述的证据表明存在一些巨大效应的基因和效应越来越小但数目不断增加的基因。那么，我们从何处停止计数对性状有影响的那个位点呢？第二个困难与连锁有关。我们当作一个位点的是染色体的一个片段，而该片段在选择过程中并没有重组。在识别连锁过程中，在这个含义上位点被称为有效因子，它们的数目被称为有效数目。尽管有这些困难，简要考虑如何估计有效因子数目还是值得的。

位点有效数目可由等式［12.1］来估计，为 $n = R_T^2/8\ \sigma_A^2$，其中 R_T 是选择上下限之差，σ_A^2 是基础群体加性方差。由等式［12.1］得到的估计值只有在下列 3 个条件下才是有效的：（1）所有有利等位基因都已固定在两个极限上；（2）所有基因都有相等的效应；（3）所有基因的起始频率为 0.5。不符合条件（1）和（2）导致 n 的估计值过低。不符合条件（3）导致 n 被过高估计，因为那时 σ_A^2 将低于要求的基因频率为 0.5 时的值。不是在基础群体中而是在处于上下限的系群间杂交的 F_2 和随后世代中估计 σ_A^2 可以满足条件（3）的要求。于是不同系群的所有基因的频率都为 0.5。但是连锁的问题随之扩大了，因为所有相关基因在 F_1 代处于完全连锁不平衡状态。随着因重组导致的不平衡减少，在 F_2 代以后的世代中遗传方差也将逐渐减小，有效因子数目则相应增加。如果来自 F_1、F_2 和回交世代的数据是有效的，则能够克服条件不符合产生的一些困难（Lande，1981）。

对 4 个实验由等式［12.1］估计的数字列于表 12-2。它们从 2 变化到 98。窝产仔数的估计值 2 显得太低以致不能被人相信，这表明作出的假设有严重错误。一个相似的实验，但只是向上选择，按照所有遗传变异是由起始基因频率为 0.25 的隐性基因引起的假设，却产生 $n = 164$ 的一个估计值（Eklund 和 Bradford，1977）。依据同样的假设，对表 12-2 中实验的上选反应进行分析，得出 $n = 25$。根据所作假设，这个巨大差异强化了这些基因数目估计值的可疑性。

依据同样的条件和限制，从有效因子数目中获得基因的标准化效应是可能的。重新安排先前给出的加性方差的表达式，即 $\sigma_A^2 = \frac{1}{2}na^2$，导致基因的标准化效应为 $2a/\sigma_P = 2h\sqrt{2/n}$，其中 h 是遗传力的平方根。在表 12-2 的 3 个实验中得到的数值接近于 0.2；窝产仔数实验对 $n = 2$ 给出一个值 0.95，或对 $n = 25$ 给出 0.21。可从等式［11.8］中近似估计出对单个基因起作用的选择系数 s。选择强度 i 非常接近 1，因此 s 值大致为 0.2。标记基因的使用使得定位对选择反应有贡献的一些基因成为可能。将在第 21 章中更详细全面地考虑多少个位点引起数量性状变异的问题。

习　题

12.1　从下列实验的结果计算表 12-2 中列出的数量（最后两个数量必须以猜测为基础）。提高小鼠3～6周龄增重的选择进行了 43 个世代。没有减少增重的选择，但有一个未选择对照。开始的 19 个世代，反应是线性的，大约在第 34 世代达到极限。需要的数据都在表中，雌雄平均。括号中的数字是数据涉及的世代数。

现实遗传力（1～19）	0.20
表型标准差（1～19）	2.10g
每世代平均选择差（1～19）	2.25g
选择系群的平均数（34～43）	23.55g
对照系群的平均数（34～43）	11.90g
有效群体大小	33

数据来自 Barria，N，& Bradford，G.E.（1981）*J. Anim. Sci.*，**52**，729～738。

［解答序号：67］

第 13 章　选择：Ⅲ．来自亲属的信息

就选择而论，直到现在我们都假定为选择性状而对个体进行度量，并且根据个体表型值选择最佳个体作为亲本。可是，个体本身的表型值并不是其育种值信息的唯一来源；亲属的尤其是全同胞或半同胞的表型值也提供附加的信息。的确，对一些性状，亲属的数值提供了唯一可用的信息。作为一个明显的例子，泌乳量不能在公畜上度量，因此公畜的育种值只能根据其雌性亲属的表型值来判断。

在对动物育种应用选择中利用亲属信息十分重要，这有两个原因。首先，被选择的性状经常是低遗传力的，对这些性状来说，多个亲属的平均值与个体本身的表型值相比，经常对育种值提供一种更可靠的指导。其次，当选择的结果在经济上有收益时，则即使反应有很小进展也足以补偿因采用最佳方法所付出的额外努力。本章我们将概述关于利用亲属信息以及挑选最佳选择方法的原理。

如果对群体的家系结构加以考虑，那么我们能计算各个家系的平均表型值；这被称为*家系均值*。于是，假定我们有一个群体，其中个体按家系分组，它们可能是全同胞或半同胞，而且我们有每一个体和每一个家系平均值的度量值。那么如何利用来自家系平均值的附加信息？这个问题最好通过一个特殊的例子来说明。表 13-1 给出了一些假定的但符合实际值的小鼠窝产仔数。有 16 个个体，它们的表型值列在表的主体部分。这些个体被分成 4 个全同胞家系：A、B、C、D，每一个家系包括 4 个个体。我们要从这 16 个个体中选出最好的 4 个。根据个体表型值选择，我们能够毫无困难地选出个体 A1、B1 和 A2 其值分别为 13、11 和 10。但是现在值为 9 的有两个个体，B2 在良好的家系内，D1 则在较差的家系内。我们该选哪一个呢？决断依赖于家系间的差异主要是遗传的还是环境的。如果这些差异是遗传的，我们就要选 B2，因为较好的家系均值表明它有一个较好的育种值。另一方面，如果家系间的差异主要是环境的，我们将选 D1，因为较低的家系均值表明它所处的环境恶劣，尽管有此不利因素，它仍然表现良好。问题不仅在于区别同一表型值的不同个体，而且要紧的是要找出家系均值的适当权重。借助正确的加权，可令我们选值为 8 的 A3 来代替值为 9 的 B2。应用被提出的原理会证明，如果这些值是小鼠的窝产仔数，那么实际上这将是最好的方法（见例 13.1）。

为了计算家系均值的最佳权重，只需要了解三方面的情况：家系种类（是全同胞还是半同胞）、家系内个体数目（即家系大小）、各家系成员间这个性状的表型相关。看来解决复杂问题需要的信息竟如此令人惊奇的简单；但对所用原理的解释并不如此简单。将从两个方面给出解释。首先我们要将遗传力的概念扩展，作为选择反应的决定因素。这里没有引入新的原理并且使上面提出问题容易得到公正的解决；但用它来解决实践中发现的较复杂问题却不方便。可是，在“指数选择”的标题下，我们将简洁地阐明更一般的解决方法。这允许将来自不同种类亲属，例如来自亲本的和同胞的信息结合起来。采用第 19 章将阐述的方法，它还允许将来自相关性状的信息用于辅助选择。

表 13-1 用于选择的个体值和家系均值的实例

个体	家系			
	A	B	C	D
1	13	11	7	9
2	10	9	7	5
3	8	6	6	3
4	5	6	4	3
家系均值	9	8	6	5
全群均值		7		

选择标准

作为群体均值离差的个体表型值的度量 P 是两部分之和：其家系均值与群体均值的离差 P_f 以及个体与家系均值的离差 P_w（家系内离差）；因此：

$$P = P_f + P_w \qquad [13.1]$$

于是，根据对这两部分给以的不同重视或不同的权重，则有多种选择方法。下面能列出 3 种简单的选择方法。第一，我们可以只根据个体值进行选择，像前两章假定的那样，即对两个组分 P_f 和 P_w 给以相等的权重，这叫做个体选择。第二，我们可以只根据家系均值 P_f 进行选择，即给家系内离差 P_w 以 0 权重，这叫做家系选择。应用于表 13-1，家系 A 的所有 4 个个体将被选留。第三，我们可以单独根据家系内离差 Pw 进行选择，而给家系均值 P_f 以 0 权重，这叫做家系内选择。应用于表 13-1，4 个家系每一家系内的最好个体将被选留。

不用以上 3 种简单方法，我们可以同时考虑 P_f 和 P_w 两个组分，并给它们以不同的权重，权重的选取要求充分利用信息的两个来源。这就叫做最佳合并选择或合并选择，或更一般地，称为指数选择。它代表获得最大反应速度的一般方法，另外那 3 种较简单的方法是两种信息来源给以 0 或 1 权重的特殊情形。因此理论上指数选择法总是最好的方法。P_f 和 P_w 的适宜权重将在以后阐明。

简单方法

3 种较简单方法的突出特征如下。

个体选择 仅仅根据个体本身的表型值来对个体选择。通常，这种方法应用起来最为简单，在许多情况下，它产生最快的反应。因此除非有充足的理由来选用另一种方法，否则应当使用个体选择。群体选择经常作为个体选择的一个术语，尤其是被选个体集中在一起进行交配时，例如像瓶中的果蝇。当交配可控制或可记录时，如小鼠或较大动物，术语“个体选择”才更专化性地使用。

家系选择 根据家系的表型平均值，整个家系作为一个单位被选择或被淘汰。因此，个体值除决定家系均值外没有其他作用。换句话说，给家系内离差以 0 权重。家系可以是全同胞或半同胞，亲缘关系较远的家系几乎没有什么实际意义。

当所选性状具有低遗传力时是最适宜采用家系选择的主要情形。家系选择的效果取决

于个体的环境离差在家系均值中倾向于互相抵消这一事实。因此，家系的表型均值可非常近似地作为其基因型均值的一个度量，当环境离差在表型方差中占很大组分时，或者换句话说当遗传力低时，获得的进展就更大。另一方面，家系成员共同的环境变异削弱了家系选择的效果。如果共同环境组分大，它就会掩盖家系间的遗传差异，家系选择将相应地无效。家系选择效果中另一个重要因素是家系中个体的数量，或家系大小。家系越大，表型均值与基因型均值之间的一致性就越紧密。因此，有利于家系选择的条件是低遗传力、共同环境引起的变异极小以及大的家系。

家系选择应用过程中有实际困难，特别是在实验群体中。这些困难是由选择强度和避免近交之间的对立引起的。保持尽可能低的近交率一般来说是理想的。考虑到近交，如果最小数目的亲本被固定，比如说 10 对，那么在家系选择下，10 个家系必须都入选，这是因为每一家系仅代表前一世代的一对亲本。如果选择强度要达到一合理的高度，则繁育和度量的家系数量很可能是此数量的 2 ~ 4 倍。因此，家系选择要占用大量的空间，如果繁育空间有限，则家系选择下所能达到的选择强度可能很小。下面两种方法是家系选择的变形。

同胞选择　正如我们已注意到的，一些性状不能在亲本个体上度量，选择只能基于亲属的表型值进行。这等于家系选择，但又有不同：此时被选个体对估计家系均值不起作用。这种差异影响受家系大小控制的反应方式。在何处作出区分是重要的，当所选个体未被度量时我们将使用术语*同胞选择*，当被选个体被度量并包括在家系均值中时使用家系选择。当家系很大时，这两种方法是等价的，于是，术语家系选择被理解为含有上述两种意思。

后裔测验　后裔测验是动物育种中广泛应用的一种选择方法。我们对它不作详细讨论，只把它当做家系选择的一种形式来处理。正如名字所暗示的，选择标准是一个个体后裔的平均值。初看起来，这似乎是理想的选择方法并最容易评估，因为就像我们在第 7 章所看到的一样，个体后代的平均值十分接近我们所能得到的个体育种值的直接度量，事实上它就是育种值在实践上的定义。然而，在实践中，后裔测验有较长世代间隔的严重缺点，因为直到后裔都已度量才能完成对亲本的选择。因为不可避免的世代重叠以及因为有可能出现应选择哪一世代，是亲本还是后裔的意义含糊的问题，所以，用后裔测验来作选择的评价很容易发生混淆。后裔均值是用来判断亲本的，只有当亲本已测验完并等候选择时，后裔才准备用作亲本。这样，中选亲本及其后裔二者同时用作亲本。通过把后裔测验当作家系选择的修改形式可以部分地克服解释的困难。后裔构成家系，通常是半同胞家系，选择就是按照上面描述的方式根据家系均值在它们之间进行。唯一的不同是中选家系因允许其亲本继续繁殖而增加了大小。另外，家系中年龄较小的成员不参与家系均值估计，因此是用同胞选择法来选择它们。通过未度量个体来增加中选家系的大小并不会提高选择准确度，但是它降低了更新速度因而提高了能达到的选择强度。这是后裔测验的主要优点，但是这种优点只在大规模的实施过程中当空间的限制不引起近交的危险时才能实现。

家系内选择　选择标准是每一个体与其所属家系均值的离差，超出它们家系均值最大数量的那些个体被认为是最理想的。这和家系选择相反，家系均值给以 0 权重。家系内选择优于其他方法的主要条件是家系成员具有很大的共同环境方差组分。猪或小鼠断奶前的

生长可作为这种性状的实例引用。个体断奶体重的大部分变异可归因于母体，因此对家系成员是共同的。家系内选择将从选择控制的变异中消除这个大的非遗传组分。家系内选择的一个重要的实际优点，特别在实验室实验中，是它节约繁殖空间，同样的原因，家系选择在空间上是耗费很大的。如果要进行单对交配，则每一家系必须选出两个成员以代替其亲本。这意味着各家系等量地为下一代提供亲本，我们在第 4 章看到的系统致使有效群体大小是实际数量的两倍。因此当实行家系内选择时，保持近交率低于某一确定值所需要的繁殖空间只是个体选择下所需要的一半。

反应的预测

为了评价不同选择方法的相对优劣，我们必须推导每一种方法的期望反应。在第 11 章中已阐明的个体选择这里不再附加说明。在等式［11.3］中给出期望反应，为 $R = i\,\sigma_P\,h^2$，其中 i 是选择强度（即以标准差表示的选择差），σ_P 是标准差，h^2 是个体表型值的遗传力。家系选择或家系内选择下的期望反应可用类似方法获得。在家系选择下，选择标准是家系成员的表型平均值，因此家系选择的期望反应是：

$$R_f = i\,\sigma_f\,h_f^2 \qquad [13.2]$$

其中，i 是选择强度，σ_f 是家系均值的观察标准差，h_f^2 是家系均值的遗传力。同样，家系内选择的期望反应是：

$$R_w = i\,\sigma_w\,h_w^2 \qquad [13.3]$$

其中，σ_w 是标准差，h_w^2 是家系内离差的遗传力。

遗传力 应用于家系均值或家系内离差的遗传力概念没有引入新的原理。很简单，遗传力是表型方差中加性遗传方差数量所占的比率。这些遗传力可用个体值的遗传力（我们将继续简单地称为遗传力，用符号 h^2 表示）、家系成员间的表型相关、家系内的个体数量来表示，所有这些都可以通过观察来估计。为了获得恰当的表达形式，我们需要重新考虑方差的观察组分是如何由有因组分组成的，就像第 9 章和第 10 章解释的一样（尤其要看表 9-5 和表 10-7）。首先让我们通过假设所有家系都包含巨大数量的个体来简化问题，这样所有家系的平均数就可无偏估计。首先考虑表型方差。家系内成员间的组内相关 t 是组间组分除以总方差：$t = \sigma_B^2/\sigma_T^2$。因此，组间组分可以表示为 $\sigma_B^2 = t\,\sigma_T^2$，组内组分可表示为 $\sigma_W^2 = (1-t)\sigma_T^2$。这表明表型方差可分剖为观察组分。总方差，这里写作 σ_T^2，是表型方差，在有因组分的背景下我们将写作 V_P。以同样的方式，加性方差在家系间和家系内的分剖可用育种值的相关 r 的形式表示。因而家系间加性方差是 rV_A，家系内加性方差是 $(1-r)V_A$。这些双重的分剖汇总在表13-2。

加性和表型方差的分剖同时导出了家系均值和家系内离差的遗传力，因为这些遗传力简单地就是加性方差与表型方差的比率。因此，当家系很大时，家系均值的遗传力是 rV_A/tV_P 或 $(r/t)h^2$，因为 V_A/V_P 是个体数

表 13-2 大的家系间和家系内的方差分剖

观察组分	加性方差	表型方差
家系间，σ_B^2	rV_A	tV_P
家系内，σ_W^2	$(1-r)V_A$	$(1-t)V_P$

值的遗传力 h^2。不同亲属关系的 r 值在表 9-3 中给出；全同胞 $r=\frac{1}{2}$，半同胞 $r=\frac{1}{4}$。为了在以后能同时讨论全同胞和半同胞家系，我们将在公式中保留符号 r 而不代入适当的值 $\frac{1}{2}$ 或 $\frac{1}{4}$。

前面所述家系均值和家系内离差的遗传力通过假定家系很大而简化了。实际应用中，这种简化是不合适的，现在我们必须通过考虑有限大小的家系来消除它。可是，我们还假定所有家系的大小是相等的。由于以下原因，我们必须考虑家系的个体数。如果选择基于家系均值或与家系均值的离差，那么我们涉及的是观察均值，而不是真正均值。换句话说，我们并不涉及目前已讨论过的方差观察组分，而涉及观察均值和观察的家系内离差的方差。各组的观察均值是有抽样方差的，它来自组内方差。如果一个组有 n 个个体，那么组均值的抽样方差是 $(1/n)\ \sigma^2_W$，其中 σ^2_W 是组内方差组分。这样观察的组均值方差增加了 $(1/n)\ \sigma^2_W$，而观察的组内离差方差相应减少同样的数量。因此，家系大小为 n 的观察方差由表 13-3 显示的观察组分组成。将观察组分变换成表 13-2 中的有因组分，现在就能找出进入观察方差的有因组分。它们列在表 13-3 的右边两栏。

表 13-3　家系大小为 n 的观察方差的组成

观察方差	观察组分	有因组分	
		加性	表型
家系均值的，σ^2_f	$\sigma^2_B+\frac{1}{n}\sigma^2_W$	$\frac{1+(n-1)r}{n}V_A$	$\frac{1+(n-1)t}{n}V_P$
家系内离差的，σ^2_w	$\sigma^2_W-\frac{1}{n}\sigma^2_W$	$\frac{(n-1)(1-r)}{n}V_A$	$\frac{(n-1)(1-t)}{n}V_P$

为获得家系均值和家系内离差的遗传力，我们只须用观察方差的加性组分除以表型组分即可。这样家系均值的遗传力是：

$$h^2_f=\frac{1+(n-1)r}{1+(n-1)t}h^2 \qquad [13.4]$$

家系内离差的遗传力是：

$$h^2_w=\frac{1-r}{1-t}h^2 \qquad [13.5]$$

同胞选择必须与家系选择区别开来，区分点的差异在于中选个体没有度量。采用等式 [10.2] 的方式，通过将遗传力视为一个回归可以最佳地推导出适宜的遗传力。在这种情况下，遗传力是未度量个体的育种值对其被度量同胞的表型均值的回归。它们的协方差简单地就是全同胞或半同胞协方差，即 rV_A，它既不受度量个体数量也不受未度量个体数量的影响，其原因已在第 9 章阐述。因此回归为 rV_A/σ^2_P，其中 σ^2_P 是表 13-3 给出的度量个体的家系均值的观察方差。代入则给出适合于同胞选择的家系均值的遗传力：

$$h^2_s=\frac{nr}{1+(n-1)t}h^2 \qquad [13.6]$$

表 13-4 列出了不同选择方法的遗传力，它们的推导现在都已作了解释。

表 13-4 不同选择方法下的遗传力和期望反应

选择方法	遗传力	期望反应
个体	h^2	$R = i\,\sigma_P h^2$
家系	$h_f^2 = h^2 \dfrac{1+(n-1)r}{1+(n-1)t}$	$R_f = i\,\sigma_P h^2 \dfrac{1+(n-1)r}{\sqrt{n[1+(n-1)t]}}$
同胞	$h_s^2 = h^2 \dfrac{nr}{1+(n-1)t}$	$R_s = i\,\sigma_P h^2 \dfrac{nr}{\sqrt{n[1+(n-1)t]}}$
家系内	$h_w^2 = h^2 \dfrac{(1-r)}{(1-t)}$	$R_w = i\,\sigma_P h^2(1-r)\sqrt{\dfrac{n-1}{n(1-t)}}$
合并	—	$R_c = i\,\sigma_P h^2\sqrt{1+\dfrac{(r-t)^2}{1-t}\times\dfrac{n-1}{1+(n-1)t}}$

i = 选择强度（按标准度量的选择差）：假定对所有方法都相等，但不必这样。

σ_P = 个体表型值的标准差。

h^2 = 个体值的遗传力。

r 对全同胞家系，$r=\frac{1}{2}$。

对半同胞家系，$r=\frac{1}{4}$。

t = 家系成员表型值的相关。

n = 家系中个体数。

期望反应 现在来推导期望反应就容易多了。以家系选择为例。期望反应由等式 [13.2] 给出：$R_f = i\,\sigma_f h_f^2$。可是，这个表达式就这样表述没有多大用处，因为它不易与其他方法进行比较。将它转换为易于与个体选择进行比较的形式将是最方便的。通过代入等式 [13.4] 给出的家系均值遗传力 h_f^2 的表达式，并用表 13-3 右边栏中的个体表型值的标准差 σ_P ($=\sqrt{V_P}$)的形式来表示观察的家系均值的标准差 σ_f，就可以做到这一点。于是，期望反应变为：

$$R_f = i\sqrt{\left[\frac{1+(n-1)t}{n}\right]} \times \sigma_P \times \left[\frac{1+(n-1)r}{1+(n-1)t}\right]h^2$$

化简为：

$$R_f = i\,\sigma_P h^2\left[\frac{1+(n-1)r}{\sqrt{n\{1+(n-1)t\}}}\right]$$

$i\,\sigma_P h^2$ 项等于个体选择下的期望反应，因此，方括号内的表达式是家系选择与个体选择比较的因子。表达式看起来很复杂，但它只包含 3 个简单的数量：n，它是家系大小；r，对于全同胞家系它是$\frac{1}{2}$，对于半同胞家系它是$\frac{1}{4}$；t，它是表型组内相关。

不同选择方法的期望反应列于表 13-4，各个表达式都用有利于与个体选择进行比较的方式来表达。

合并选择

这里只对合并选择（combined selection）作简要的论述，因为在以后将更充分地阐述它。在应用合并选择时所采用的适宜权重因子可推导如下。以前我们知道，个体表型值由两部分组成，即家系均值和家系内离差，$P = P_f + P_w$，每一部分都给出有关个体育种值的

一些信息。在第 10 章我们看到遗传力等于育种值对表型值的回归（等式［10.2］），因此，从表型值得到个体育种值的最佳估计值是 h^2P（等式［10.4］）。这个概念能独立应用于表型值的两个部分，因为这两个部分互不相关且为育种值提供独立的信息。因此考虑表型值的两个部分，个体育种值的最佳估计值可由复回归方程给出：

$$\text{期望育种值} = h_f^2 P_f + h_w^2 P_w \qquad [13.7]$$

P_f 作为与群体均值的离差而度量，P_w 则作为与家系均值的离差而度量。因此最有效利用两种信息来源的加权因子是两个遗传力，它们分别是对家系均值和家系内离差的偏回归系数。如果用表 13-4 中的两个遗传力值代入等式［13.7］，我们将看到，h^2 项是两个权重因子的共同部分，因此可以省略 h^2 这一项而不影响相对权重。这样给出的期望育种值为：

$$E(A) = \left[\frac{1-r}{1-t}\right]P_w + \left[\frac{1+(n-1)r}{1+(n-1)t}\right]P_f$$

在实践中，用个体值代替家系内离差并且将它的权重定为 1 是十分方便的。因此使用校正的家系均值，补充了个体本身提供的信息。对家系均值的适宜权重因子的整理可导出一个有价值的指数如下（Lush，1947）：

$$I = P + \left[\frac{r-t}{1-r} \times \frac{n}{1+(n-1)t}\right]P_f \qquad [13.8]$$

式中 P 是个体值，P_f 是家系均值与群体均值的离差，个体本身包括在家系均值内。注意如果 t 大于 r，则 P_f 的权重是负的。当相关中有很大的环境组分时，这种情况才会出现；于是家系均值是环境而不是育种值的标志。以适宜于和个体选择比较的形式表示的合并选择的期望反应给在表 13-4 中。它的由来见 Lush（1947）。

例 13.1

通过将表 13-1 中的数字应用于等式［13.8］来说明合并选择的运作，这些数字是小鼠窝产仔数的实际值。这里再次把表型值以及作为与全群均值离差的家系均值列于表 13-5a。窝产仔数的全同胞相关大约是 $t = 0.1$。用 $r = 0.5$、$n = 4$ 还有 t 代入等式［13.8］得到有价值的指数为：$I = P + 2.46P_f$。对每一个体这样计算得来的指数列于表 13-5b。由合并选择选出的个体和由个体选择选出的个体之间唯一的差异是排名第 4 的个体，合并选择选 A3 而不是 B2 或 D1。

表 13-5a　表型值和家系均值

	A	B	C	D
1	13	11	7	9
2	10	9	7	5
3	8	6	6	3
4	5	6	4	3
P_f	+2	+1	−1	−2
全群均值		7		

表 13-5b 表型值和家系均值

	A	B	C	D
1	17.9	13.5	4.5	4.1
2	14.9	11.5	4.5	0.1
3	12.9	8.5	3.5	-1.9
4	9.9	8.5	1.5	-1.9
全群均值		7		

各种方法的相对优点

任一选择方法相对于任何其他选择方法的优点可由表 13-4 给出的期望反应来揭示。可是，表示相对优点的公式是很繁琐的，以一种复杂的方式取决于表型相关 t、家系大小 n 以及家系是全同胞还是半同胞即 r。相对优点以图解方式汇总在图 13-1 中。它揭示了 3 种简单方法相对于合并选择的期望反应，合并选择一定总是最好的方法。用两个极端家系大小的相对反应对表型相关 t 绘制成图：（a）家系很大，（b）家系大小为 2。假定选择强度 i 对所有方法都相同。虽然图解中的结论不能直接应用于实际育种中，但为说明其理由，还是将它们列出如下。第一，将 3 种简单方法与合并选择比较，简单方法的期望反应低于合并选择的，很少多于 10%，从不大于 20%；第二，比较 3 种简单方法，在很大的 t 值范围内个体选择都是最好的。其原因是个体选择利用了整个的加性遗传方差，而家系选择仅利用了家系均值之间的方差，家系内选择仅利用了家系内方差。当表型相关 t 低时，家系选择比个体选择好。低的同胞相关暗示遗传力低和几乎没有来自共同环境的相似性。一般来说，就是这些条件使得家系选择比个体选择好。当同胞相关很高时，家系内选择比个体选择好。很高的相关只能由巨大的共同环境方差组分 V_{Ec} 引起；这是家系内选择优于个体选择的条件，但这种情形极少发生。可是，还要考虑其他因素。第一个就是近交。家系选择似乎导致中选的亲本来自较少的几个家系，除非选择强度相应地降低。结果家系选择的近交率似乎比个体选择的大。确实，家系均值的任何使用，比如合并选择，都倾向于增加近交率。相反，家系内选择似乎降低近交率。第二个因素是第 11 章中解释过的加性方差的减少。正如表 11-5 所示，它是由于选择产生的不平衡而降低的唯一的家系间方差。因而家系内选择优于家系间选择之处正是它能利用了较大数量的加性方差。第三个因素是经济方面的，它使得图 13-1 中的比较结果不能直接与实际育种工作相关联。同时获得个体和家系记录的费用可能使合并选择不经济或不实际。在动物育种中，个体记录容易获得，但获得合并选择需要的谱系记录可能是昂贵的。可在较少有效的个体选择和较昂贵的合并选择之间进行挑选。另一方面，在植物育种中，家系可能容易鉴别出来，而单株植物记录可能难于获得，或者它们没有意义。这种情形下家系选择可能是唯一可行的方法。应用于植物的家系选择的详细论述见 England（1977）。

例 13.2

为举例说明如何比较不同的选择方法，将用小鼠的 3 个性状来计算它们的相对优点。假定家系是单一窝次。断奶重具有很大的共同环境组分，它造成的全同胞相关高达 0.8。6 周龄体重的共同环境组分较小，但仍很大，它造成的相关为 0.6（见例 10.5）。窝产仔

数，作为成年母鼠的一个性状，遗传力低且共同环境效应小，它造成的全同胞相关很低为 0.1。假定所有选择方法的选择强度都相同，虽然这在实践中是不可能的。假定家系大小为 $n=4$。家系是全同胞，因此 $r=0.5$。通过代入这些 n、r 值和适当的 t 值，从表 13-4 中计算了相对于个体选择的期望反应。相对反应给在表 13-6。对断奶重和窝产仔数，合并选择比个体选择大约好 20%，但对于 6 周龄体重，合并选择实际上没有优势。对于窝产仔数，家系选择比个体选择要好 10%（如果公鼠也要选择，不管应用于母鼠的是什么方法，这些公鼠必须通过同胞选择来选留）。对这 3 个性状中的任一性状来说，家系内选择并不比个体选择好，但是对于断奶重，若家系大小大于 5，则不然。

表 13-6　不同选择方法对性状的影响

	断奶重	6 周龄体重	窝产仔数
t	0.8	0.6	0.1
等式[13.8]中 P_f 的权重	−0.71	−0.29	+2.46
合并：R_c/R	1.18	1.01	1.19
家系：R_f/R	0.68	0.75	1.10
家系内：R_w/R	0.97	0.68	0.46

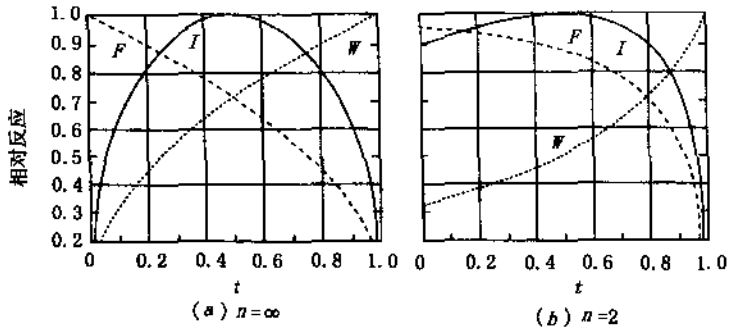


图 13-1　对全同胞家系不同选择方法的相对优点。用相对于合并选择的反应对表型组内相关 t 绘图
（I = 个体选择；F = 家系选择；W = 家系内选择）

例 13.3

已用果蝇（*Drosophila*）做实验进行同胞选择与个体选择的比较（Clayton 和 Robertson，1957）。同时用全同胞和半同胞两种家系进行同胞选择。反应与具有例 11.2 给出的选择强度 $i=1.40$ 的个体选择进行了比较。表 13-7 给出了表 13-4 的公式计算相对于个体选择的期望反应所需要的数据。我们将看到半同胞的 r 值比 0.25 稍大一些。这是因为与同一雄体交配的雌体彼此间并不是完全地互不相关。期望和观察到的相对反应给在表的右边部分。预期同胞选择应当比个体选择差，这已被证明。然而在向上和向下的反应间存在一些差异，其原因不明。

表 13-7 同胞选择与个体选择的比较

	数 据		相对反应，R_s/R		
	全同胞	半同胞		全同胞	半同胞
i	1.33	1.27	期望值	0.832	0.614
n	12	20	观察值，向上	0.618	0.527
r	0.50	0.275	观察值，向下	0.919	0.635
t	0.265	0.121			

指数选择

选择的最优方法是使用个体育种值的所有可利用的信息，并把这些信息综合为一个有价值的指数。当存在两个以上的信息来源时，上面所给出的合并家系均值与家系内离差的解决方法不能很好地适用于较复杂情况。例如，可能有来源于个体、双亲、全同胞、半同胞、其他亲属的信息。或者，如果是限性性状，则个体本身不能度量的信息只能来自于不同种类的亲属。因此，我们的目的是将所有信息合并成一个指数，并据此对个体进行选择。不使用矩阵方法，指数是不容易构建的，尤其如果有两个以上的信息来源时。技术上的细节超出了本书范围，这里仅给出有关原理的简要叙述。更详细的阐述见 Nordskog (1978)、Lin (1978)、Nicholas (1987) 和 Van Vleck (1993)。

指数的构建

指数是个体育种值的最佳线性预测，它采取育种值对所有信息来源的复回归形式。首先考虑最简单的情况，即我们拥有的唯一信息是作为与群体均值离差的个体本身的表型值 P。于是预测或期望的育种值为 $E(A) = b_{AP}P$，其中 b_{AP} 是育种值对表型值的回归。在这种情况下，$b_{AP} = h^2$ (等式 [10.2])。现在假定有若干种信息 P_1、P_2、P_3 等，其中每一个 P 是个体或一组亲属的表型值。这些以表型值形式出现的信息将被称为“度量值”；它们将作为与群体均值的离差来表达。于是，个体的指数为：

$$I = b_1P_1 + b_2P_2 + b_3P_3 + \cdots \qquad [13.9]$$

其中，b 是对各度量值进行加权的因子。问题是要找出每个加权因子的最佳值。这可以通过找出使指数与育种值间相关 r_{IA} 达到最大的 b 值来做到。使 r_{IA} 最大等于使指数值与 I 对 A 的线性回归的离差平方和即 $\sum(I - A)^2$ 最小。于是，由此产生的 b 值是个体育种值对各度量值的偏回归系数。r_{IA} 的最大化是计算偏回归的标准方法，这里不必再作解释。求解最大值可导出一组联立方程式，方程的个数与度量值个数一样多，这些方程的解给出等式 [13.9] 中要使用的 b 值。待解的这些指数方程在下面给出 (等式 [13.10])。给出了 3 个度量值的方程；很容易看到，这些方程对较多或较少的度量值如何进行扩展或减少。每一个方程将在左边的度量值的表型方差和协方差与在右边的度量个体的加性遗传方差和协方差联系起来。符号简要叙述如下。P 表示由下标数字表示的度量值的表型方差或协方差。例如，P_{11} 是度量值 1 的表型方差，P_{12} 是度量值 1 和 2 的表型协方差。育种值的方差和协方差类似地记作 A。方程为：

$$\left.\begin{aligned} b_1P_{11} + b_2P_{12} + b_3P_{13} &= A_{11} \\ b_1P_{21} + b_2P_{22} + b_3P_{23} &= A_{21} \\ b_1P_{31} + b_2P_{32} + b_3P_{33} &= A_{31} \end{aligned}\right\} \qquad [13.10]$$

当然，要解这些方程必须代入 P 和 A 的数值。P 和 A 可以用下面的参数来表示：表型方差 V_P，这里记为 σ^2；个体值的遗传力 h^2；个体间的表型相关 t；表 9-3 中列出的亲缘关系 r。当“度量值”是一群个体的均值时，则群体中个体数 n 也是必需的。解这些方程有标准的计算机程序。将通过简化为只考虑两个度量值的特殊例子来说明指数的获得。Becker（1984）阐述了其他一些例子。

个体和一个亲属　设度量值 1 是要计算指数的个体的，度量值 2 是它一个亲属的。那么度量值的表型方差相同，即 $P_{11} = P_{22} = \sigma^2$。表型协方差是 $P_{12} = P_{21} = t\sigma^2$。加性方差是 $A_{11} = h^2\sigma^2$，加性协方差是 $A_{21} = rh^2\sigma^2$，其中 r 是亲属和个体之间的亲缘系数。各部分被 σ^2 除后，待解方程变为：

$$b_1 + tb_2 = h^2$$
$$tb_1 + b_2 = h^2$$

化简之后，方程的解为：

$$b_1 = \frac{h^2(1-rt)}{1-t^2};\ b_2 = \frac{h^2(r-t)}{1-t^2}$$

将 b_1 和 b_2 的值代入等式［13.9］即得指数。

为了选择个体的实用目的，再度量（rescale）指数可能是方便的，这样个体本身的度量值给以 1 的权重。于是给予 P_2 的权重是 $b_2/b_1 = (r-t)/(1-rt)$，再度量指数为：

$$I' = P_1 + \left(\frac{r-t}{1-rt}\right)P_2$$

对于这种形式的指数，P_1 可用实际值，但 P_2 必须是群体均值的离差；于是指数值是以度量值为单位的。最初形式的指数是预测个体育种值的；再度量形式的指数则是经亲属信息校正后的表型值。任何指数都可作类似再度量。可是必须注意，如果要预测指数选择的反应，则必须考虑到对指数的任何再度量。反应的预测将在以后解释。

母亲和一个父系半姐妹　这是例证按雌性性状对公畜选择，像产奶量或产蛋量。要计算指数的个体不能度量。为了使指数方程与等式［13.10］的方程可比，我们可以认为度量值 1 缺省，度量值 2 是母亲的，度量值 3 是半姐妹的。因此等式［13.10］的相关部分为：

$$b_2P_{22} + b_3P_{23} = A_{21}$$
$$b_2P_{32} + b_3P_{33} = A_{31}$$

表型方差还是相等的，$P_{22} = P_{33} = \sigma^2$。假定母亲和半姐妹没有亲缘相关也没有环境相关，因此 $P_{23} = P_{32} = 0$。A_{21}是个体与其母亲的加性协方差，它等于$\frac{1}{2}h^2\sigma^2$。A_{31}是个体与其半同胞的加性协方差，它等于$\frac{1}{4}h^2\sigma^2$。于是将指数方程直接求解：

$$b_2 = \frac{1}{2}h^2;\ b_3 = \frac{1}{4}h^2$$

给出指数为：

$$I = \frac{1}{2}h^2P_1 + \frac{1}{4}h^2P_2$$

个体和同胞均值　这是本章开头介绍问题的那种情况，前面描述为“合并选择”。像前面一样，个体仍是度量值 1，因此 $P_{11} = \sigma^2$，$A_{11} = h^2\sigma^2$。家系均值的方差给在表 13-3，从中可知 $P_{22} = [1 + (n-1)t]\ \sigma^2/n$。通常个体包括在家系均值内是方便的。于是个体与家系均值的协方差等于家系均值方差。这样 $P_{12} = P_{21} = P_{22}$。同样，加性协方差等于家系均值的加性方差。从表 13-3 中获得加性方差，给出 $A_{21} = [1 + (n-1)r]h^2\ \sigma^2/n$。因此指数方程为：

$$b_1 + Kb_2 = h^2$$

$$Kb_1 + Kb_2 = kh^2$$

其中，$K = \dfrac{1 + (n-1)t}{n}$，$k = \dfrac{1 + (n-1)r}{n}$

解出 b_1 和 b_2 即得指数：

$$I = \frac{h^2(1-k)}{(1-K)}P_1 + \frac{h^2(k-K)}{K(1-K)}P_2$$

如果不同个体的同胞数不同，对每一个体必须分别计算 k 和 K。如果同胞数是常数，则指数可再度量为与等式［13.8］相同的形式。

从同胞获得的信息的有用性取决于家系大小 n。因为通常半同胞比全同胞多得多，半同胞家系比全同胞家系提供的信息要多，尽管其遗传亲缘关系不甚紧密。Osborne（1957）已检查过应用家禽指数的相对优势。

包含均值的其他“度量值”　在其他情形中应用上面由 K 和 k 表示的数量时，有两点需要注意。首先，一组亲属彼此间的相关方式不同于它们与个体相关的方式。于是必须注意使用适宜的 t 值和 r 值。例如，P_2 可能是一组全同胞的均值，而对个体而言，全同胞又变为半同胞。于是适合于 P_{22} 的 t 值是全同胞的相关，但适合于 P_{12} 和 P_{21} 的 t 值以及适合于 A_{21} 的 r 值则是半同胞的相关。其次，个体“度量值”可能是几个重复记录的均值。于是，必须使用重复力来替代 t。

准确度

指数值与育种值间的相关 r_{IA}，在指数的构建中它达到最大值，叫做指数的准确度。准确度提供了比较指数或任何选择标准的一种方便的方法，因为相关越高，作为育种值预测的标准就越好。对于个体选择，标准简单地就是个体的表型值。因此个体选择的准确度是表型值与育种值的相关，它是 h，即遗传力的平方根（等式［10.3］）。指数准确度 r_{IA} 计算如下。

首先，我们需要知道指数值的方差。由等式［13.9］可知：

$$\sigma_I^2 = b_1^2P_{11} + b_2^2P_{22} + \cdots + 2b_1b_2P_{12} + \cdots \qquad [13.11]$$

其中方差和协方差以等式［13.10］的符号表示。这种表达式可以合并为易于计算的形式。重新整理各项，有：

$$\sigma_I^2 = b_1(b_1P_{11} + b_2P_{12} + \cdots) + b_2(b_1P_{21} + b_2P_{22} + \cdots) + \cdots$$

用等式［13.10］代替括号中的式子得：

$$\sigma_I^2 = b_1A_{11} + b_2A_{21} + b_3A_{31} + \cdots \qquad [13.12]$$

下一步我们需要知道指数值与育种值的协方差。由下面的理由可知它就是 $\mathrm{cov}_{IA} = \sigma_I^2$。权重因子 b 是偏回归系数，这样构建的指数导致育种值对指数值的回归等于 1，即 $b_{AI} = 1$。从另一种角度讲，这就是：构建指数导致 1 单位指数值等于 1 单位预测育种值。现在有 $\mathrm{cov}_{AI}/\sigma_I^2 = b_{AI} = 1$，由此可知 $\mathrm{cov}_{AI} = \sigma_I^2$。这样，假定指数未经再度量，则相关由下式给出：

$$r_{IA} = \frac{\mathrm{cov}_{IA}}{\sigma_I\sigma_A} = \frac{\sigma_I}{\sigma_A} \qquad [13.13]$$

这里 σ_I 从等式［13.12］中获得，σ_A 是群体加性遗传方差的平方根。

相关 r_{IA} 是复相关，它的平方表示指数中合并的度量值占加性方差的部分，即 $r_{IA}^2 = \sigma_I^2/\sigma_A^2$。相反，$(1 - r_{IA}^2)$ 则是由指数中没计入的加性方差部分。这表明通过包含附加度量值的指数改良还有多大余地。

选择反应

选择反应是中选亲本的平均育种值，它可由育种值对指数值的回归来预测，即 $R = b_{AI}S$，其中 S 是指数值的选择差。代入 $b_{AI} = r_{IA}\sigma_A/\sigma_I$ 和 $S = i\sigma_I$ 得预测反应为：

$$R = ir_{IA}\sigma_A \qquad [13.14]$$

这个预测公式适用于任何指数，只要 r_{IA} 指的是实际使用的指数，就不管是怎样构建的以及是否是最优的。如果指数未作再度量，用等式［13.13］代换 r_{IA} 得：

$$R = i\sigma_I \qquad [13.15]$$

与简单的个体选择相比，很显然，指数选择应用起来要费力和昂贵得多。通过它能增益多少？将等式［13.14］中指数选择的期望反应与等式［11.4］中个体选择的期望反应（$R = ih\sigma_A$）进行比较，就能很容易地看出相对优势。只要选择强度相同，从指数选择中预期的改进相对于个体选择的改进就是比率 r_{IA}/h，这是它们各自准确度的比率。

实践　在实际动物育种中，育种值的估计存在大量的复杂情况：估计育种值的个体存在相关，一些个体有重复度量，其他个体仅有单个记录，最严重的情形是，个体组群之间存在系统环境差异。对后一种情况给出一个实例，乳用母牛饲养在管理实践不同的牛群中，这样，对公牛产奶量育种值的估计，必须按其女儿的牛群效应来校正。如果所有有关的固定环境因子的估计值已知，那么能够校正每一个体的度量值，采用校正的值可以构建选择指数。可是，校正因子的估计值通常事先是不知道的。被称为最佳线性无偏预测（BLUP）的统计方法用来同时估计固定效应和育种值。因此，传统的选择指数是具有已知固定效应的 BLUP 的一种特殊情形。只要实施选择的数据包括在分析中，BLUP 还能适应

其他复杂情况，诸如非随机交配和由个体选择引起的偏差。BLUP 假设遗传和表型方差已知。如果这不是真实的，则采用 REML 可从数据中同时估计出遗传参数。被称为（个体）动物模型的 BLUP 的一个版本被广泛用来评定动物育种数据。动物模型阐述了被度量的每一个体的育种值，解释了育种值要被估计的个体之间的所有遗传关系。因为模型中包括世代间个体间的关系，所以动物模型能够解释遗传平均值和方差的变化，因而是分析选择实验数据的最佳方法。完美的解释和数字例子见 Nicholas（1987）和 Van Vleck（1993），更多的细节见 Henderson（1977，1988），Thompson（1979，1989）以及 Kennedy 和 Sorensen（1988）。

将和需预测性状相关的性状（且这些性状本身就是选择目标）的数据合并使用，则可以改进预测育种值的指数。关于指数将在第 19 章阐述。

实际成就

在结束选择题目之前，我们应当考虑近些年来通过选择的应用家畜实际上获得了多少改良。Smith（1984，1988）已综述了这个问题。已作过遗传改良评估的性状有：家禽中肉仔鸡增重和蛋鸡产蛋数、猪和绵羊的体重和窝产仔数性状、牛的体重和产奶量。所有这些都表现出改进。用平均值的百分率表示，改良速度大都在每年1% ~2% 的范围内（不是每一世代）。可是，肉用禽类表现出 6% 的改进，肉牛只有 0.3%。除肉用仔鸡外，这些改良速度有点低于可能达到的水平。可能的改良速度由已知的遗传力、选择方法的准确度和选择强度来计算。可能的改良速度在每年1.4% ~3.2% 间变化。结论是毫无疑问还能获得相当可观的进展，而且通过改进方法来提高改良速度应当是可能的。

习 题

13.1 对具有下列参数的性状，计算家系均值的遗传力 h_f^2 和家系内离差的遗传力 h_w^2。假定选择强度相等，计算每一种情形中家系选择和家系内选择相对于个体选择反应的期望反应，即 R_f/R、R_w/R。

个体遗传力	家系类型	同胞相关	家系大小
(1) 0.1	半同胞	0.025	10
(2) 0.1	半同胞	0.025	20
(3) 0.1	全同胞	0.5	4
(4) 0.2	全同胞	0.8	4
(5) 0.2	全同胞	0.8	8

[解答序号：77]

13.2 英国大白猪日增重的半同胞相关是 0.10，全同胞相关是 0.36。当选择基于下列条件时，比较相对于个体选择进展的期望反应率。

(1) 全部来自不同母亲的 5 个半同胞的均值，中选个体是其中之一。

(2) 包括中选个体的 5 个全同胞的均值。

（3）不包括中选个体的 5 个全同胞的均值。

（4）个体与包括个体本身的 5 个全同胞的均值的离差。

数据来自 Smith，C. *et al*.（1962）*Anim. Prod*.，**4**，128 ~ 143。

［解答序号：87］

13.3　根据个体的增重和 5 个全同胞的家系均值，个体包括在家系均值内，来选择猪的日增重，则适宜的指数是什么？假定全同胞相关和习题 13.2 中的相同都为 0.36。

［解答序号：97］

13.4　如果下表数字是来自不同全同胞家系的 4 个个体猪的日增重，根据习题 13.3 解出的指数并对同胞数做一些必要的修改，则它们的优先顺序是什么？

猪	增　重		
	个体	家系均值	家系大小
A	1.6	1.3	5
B	1.5	1.6	5
C	1.5	1.6	8
D	1.3	1.7	8

群体平均值 = 1.5

［解答序号：107］

13.5　如果利用习题 13.3 中计算出的指数来选择猪的日增重，那么期望反应同个体选择期望反应进行比较，结果怎样？

［解答序号：117］

13.6　根据母亲和 10 个父系半姐妹的产奶量来构建一个选择公牛产奶量的指数。假定半姐妹全部来自不同的母亲，且它们的母亲与公牛母亲不相关。还假定半同胞之间没有环境相关。设产奶量的遗传力为 0.35（表 10-1），虽然这比大多数估计值要高。

［解答序号：127］

13.7　若根据习题 13.6 的指数从 20 头公牛中选出 1 头，且从两头母牛中根据其本身产奶量选出 1 头，预测每一世代产奶量的改进率。假定产奶量的表型标准差为 696kg（例 8.7），且假定从大量度量的群中进行选择。

［解答序号：137］

第 14 章　近交和杂交：Ⅰ. 平均值的变化

现在我们将注意力转向近交，它是育种者改变群体遗传组成的两种方法中的第二种方法。近交对繁殖力和一般活力的有害效应是育种者和生物学家所熟知的，并作为数量性状所表现的两个基本遗传现象之一已在第 6 章提及。其对立面或互补面，近交系之间或不同品种或变种之间的杂交所引起的杂种优势现象同样也是众所周知的，并成为动植物改良的一种重要手段。在杂种优势利用过程中生产用于以后杂交的系群是进行近交的两个主要目的之一。另一个目的则是生产遗传上一致性的品系，尤其是用于生物检定和各种不同领域研究的实验动物。可是，近交本身几乎普遍是有害的，正常情况下育种者或实验者试图尽可能避免近交，除非是为某些特殊目的。

在第 3 章给出的近交讨论中，根据基因频率和基因型频率的变化描述了近交的结果。这里我们必须指明基因和基因型频率的这些变化如何影响数量性状，以及它们如何解释观察到的近交和杂交的效应。本章将说明近交和杂交对平均值的影响，对方差的影响将在下一章说明。在第 16 章我们将考虑用近交和杂交作为动植物改良的一种手段。

要描述的近交效应不适用于自然自体受精的植物。因为近交是它们正常的交配方式，所以它们不可能进一步近交。可是，它们能杂交，于是它们确实经常表现出杂种优势，尽管这种优势低于异交物种的近交系杂交时的杂种优势。自体受精植物的改良，杂交是第一步，将在第 16 章摘要描述。

近交衰退

近交最显著的实际结果是与繁殖能力或生理机能相关的性状所表现的表型平均值降低，这种现象称为近交衰退（inbreeding depression）。近交衰退的一些例子给在表 14-1，从中我们可以看出什么类型的性状易于近交衰退，并且也可很粗略地看出效应的大小。从这些和很多其他研究的结果中，我们可作出如下概论：近交倾向于降低适合度。这样，构成适合度重要组分的性状诸如窝产仔数或哺乳动物的泌乳，对近交表现出下降；而那些与适合度不密切相关的性状则表现极小变化或没有变化。例如，果蝇（*Drosophila*）中，刚毛数和体重不变化（Kidwell 和 Kidwell，1966），而生殖力和存活力变化（Tantawy 和 Reeve，1956）。

在谈到一特定性状表现的近交衰退时，我们指的是许多系群的平均值的平均变化。通常发现分离的系群所表现的变化程度或大或小都有所不同，确实象我们在关于基因频率随机漂变的结果中所预测的一样。系群分化问题将在以后我们讨论方差变化时讨论。这里提到它只是为了强调这个事实，即现在要讨论的平均值变化指的是从一个基础群体中衍生的大量系群的平均值变化。正如我们先前描述的近交那样，我们不得不描绘由很多系群组成的“整个群体”。于是群体平均值指的是这整个群体的平均值，近交衰退指的是这个群体平均值的减少。现在让我们考虑近交中群体平均值变化的理论基础。

表 14-1　近交衰退的一些实例。近交系数每增加 10% 平均值的大概下降数：(1) 以绝对单位表示；(2) 作为非近交平均值的百分率；(3) 作为原始表型标准差的百分率。除下面标明的外，给出的衰退仅仅是由其性状被度量的个体近交引起的

		(1) 单位	(2) M 的%	(3) σ_P 的%
人				
10 岁时的身高（cm）；[Schull，1962]		2.0	1.6	37
智力测验得分（百分点）；[Morton，1978]		4.4	4.4	29
牛				
产奶量（kg）；[Robertson，1954]		135	3.2	17
绵羊 [Morley，1954]				
剪毛量（kg）		0.29	5.5	51
1 岁体重（kg）		1.32	3.7	36
猪 [Bereskin 等，1968]				
窝产仔数（出生的存活数）	(a)	0.24	3.1	9
154 日龄体重（kg）		2.6	4.3	15
小鼠				
窝产仔数；[Bowman 和 Falconer，1960]	(b)	0.56	7.2	23
6 周龄体重（g）；[White，1972]		0.19	0.6	7
玉米 [Cornelius 和 Dudley，1974]	(c)			
株高（cm）	(FS)	5.20	2.1	4
	(S)	5.65	2.3	5
种子产量（g/单株）	(FS)	7.92	5.6	25
	(S)	9.65	6.8	30

(a) 母本近交；仔畜非近交。

(b) 连续全同胞交配下与母本近交有关的衰退；一个世代的仔畜比母本近交更厉害。

(c) 与被度量植物的近交有关的衰退；通过连续全同胞交配（FS）或自体受精（S）来近交。J.W.Dudley 博士热心提供玉米的 σ_P 值。

首先，我们可以回忆和扩展第 3 章中的一些结论，开始假设选择不以任何方式影响基因频率的离散过程。由于整个群体的基因频率在近交时不变化，因此群体平均值的任何变化一定可归因于基因型频率的变化。近交引起纯合基因型频率增加和杂合基因型频率的下降。因此近交时群体平均值的变化一定与纯合子和杂合子之间基因型值的差异有关。现在让我们更准确地看看群体平均值是如何取决于近交程度的。

考察被细分成许多系群的一个群体，具有近交系数 F。采用下面的方法，将表 3-1 和 7-2中推理的结果合在一起可获得群体平均值的表达式。表 14-2 指出了一个两个等位基因位点的 3 种基因型及它们在整个群体中的基因型频率。这些频率来自表 3-1，$\bar{p}$ 和 $\bar{q}$ 是整个群体的基因频率。于是第三栏给出了图 7-1 中指定的基因型值。每一种基因型的值和频率乘在一起放在右边栏中，它们的和给出了这个位点对群体平均值的贡献。这样，仍然指单个位点效应，我们发现近交系数为 F 的群体平均基因型值为：

$$M_F = a(\bar{p} - \bar{q}) + 2d\bar{p}\,\bar{q}(1 - F) \qquad [14.1]$$

$$= M_0 - 2d\bar{p}\,\bar{q}F \qquad [14.2]$$

其中，M_0 是近交前的群体平均值，来自等式 [7.2]。因此由近交引起的平均值的变化为 $-2d\bar{p}\,\bar{q}F$。这表明只要 d 不为零，换句话说就是如果杂合子的值不同于纯合子的平均值，

那么，位点将对近交的均值变化起作用。虽然这只对一个位点上两个等位基因做了详细证明，但这个结论对具有两个以上等位基因的位点同样有效。因此可以得出下面的一般结论：近交时平均值的变化是与该性状有关的位点上显性的结果，并且变化的方向倾向于更为隐性的等位基因的值。显性可能是部分的或完全的，或它可能是超显性的；一个位点对平均值变化起作用的必要条件是杂合子不应当恰好处于两个纯合子之间的中间值。等式[14.2]也表明平均值变化的大小取决于基因频率。当$\bar{p}\,\bar{q}$最大时，即当$\bar{p}=\bar{q}=\frac{1}{2}$时，变化最大。因此，其他条件相同时，处于中间频率的基因要比处于高或低频率的基因对平均值的变化贡献更大。

表 14-2 两个等位基因位点的三种基因型及其在群体中的频率

基因型	频率	值	频率×值
A_1A_1	$\bar{p}^2+\bar{p}\bar{q}F$	$+a$	$\bar{p}^2a+\bar{p}\bar{q}aF$
A_1A_2	$2\bar{p}\bar{q}-2\bar{p}\bar{q}F$	d	$2\bar{p}\bar{q}d-2\bar{p}\bar{q}dF$
A_2A_2	$\bar{q}^2+\bar{p}\bar{q}F$	$-a$	$-\bar{q}^2a-\bar{p}\bar{q}aF$
		和	$=a(\bar{p}-\bar{q})+2d\bar{p}\bar{q}-2d\bar{p}\bar{q}F$
			$=a(\bar{p}-\bar{q})+2d\bar{p}\bar{q}(1-F)$

现在考虑影响性状的所有位点的联合效应。就多位点的基因型值加性联合来说，群体平均值由单独位点贡献的总和给出，这样：

$$M_F=\sum a(\bar{p}-\bar{q})+2(\sum d\bar{p}\bar{q})(1-F) \qquad [14.3]$$

$$=M_0-2F\sum d\bar{p}\bar{q} \qquad [14.4]$$

近交引起的平均值变化为$-2F\sum d\bar{p}\bar{q}$。

这些表达式表明了处于什么样的条件下一个数量性状将表现出近交引起的平均值的改变。主要的一个条件是定向显性，这意味着有关基因的显性在一个方向上占优势。如果增加性状值的基因对其减少性状值的等位基因呈显性，那么近交将引起群体平均值的减少，即向着更为隐性的等位基因的方向变化。可是，每一个位点的贡献也取决于它的基因频率，具有中间频率的那些基因对平均值的变化有最大的影响。

从等式[14.4]可获得的另一个结论是当位点加性联合时，近交引起的平均值的变化应当直接与近交系数成比例。换句话说，当对 F 作图时，平均值的变化应当是一条直线。如果位点间存在上位互作，则平均值和近交系数之间的关系不再是线性的。非线性是由双或多杂合子的互作离差引起的。双杂合子的频率下降与 F^2 成比例。因此，如果平均而言互作离差是正的即有利的，那么随着 F 增加，则平均值衰退率增加，如果它们是负的，则衰退率下降。没有其他互作形式影响直线性，并且无显性的上位本身不能引起任何近交衰退。至于上位如何影响近交衰退的细节，参见 Crow 和 Kimura（1970，p.79）。

实验上观察到的近交衰退的例子在图 14-1 和 14-2 中说明。总的来说，在这些例子中观察到的近交衰退，就 F 来说，确实倾向于呈线性，这可以作为位点间上位互作不十分重要的证据。可是，若干实际困难阻碍着从近交衰退率的观察中得出确切的结论。一个困难是，随着近交进行和繁殖效率下降，很快就不可避免地丢失一些个体和一些整个系群。

于是存活者就是一个经过选择的群体，理论期望值不再适用于它们。因此，近交衰退率的精确度量一般只能在早期阶段，近交系数达到高水平之前进行。另一个困难，尤其在哺乳动物研究中遇到的，是由母体效应引起的。母体特性存在于对近交衰退最敏感的性状之间。因此近交对受母体效应影响的另一个性状的影响是双重的：部分可归因于被测定个体的近交，部分可归因于母本的近交。当进行连续近交时，在早期世代中母本和后裔有不同的近交系数，因此被测性状和近交系数之间的关系不能用任何简单方式表达。由于这些困难，不可能从连续近交下观察到的衰退的精确形式中很容易地获得可靠结论。

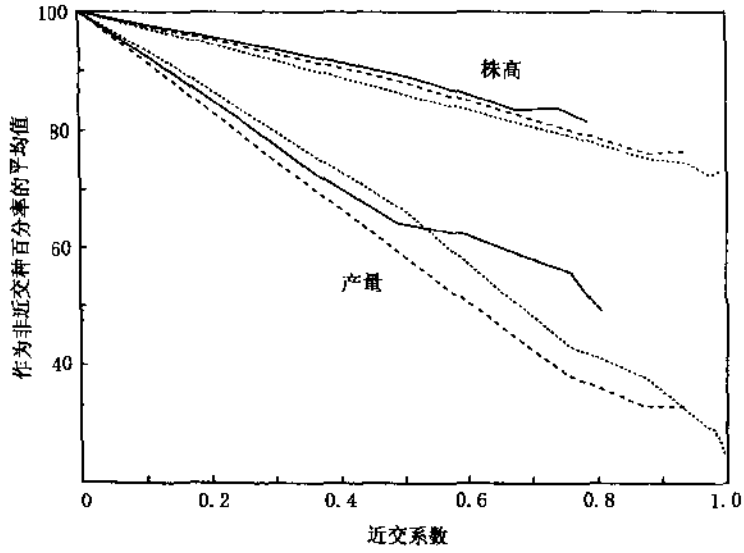


图 14-1　近交对玉米（*Zea mays*）株高和种子产量的影响

点线和短条线指连续自交；连续直线指连续的全同胞交配。不进行任何选择。

数据来自 Hallauer 和 Sears（1973）（点线），Cornelius 和 Dudley（1974）（连续直线和短条线）。

选择的效应

近交过程中对选择的忽略是一个不现实的省略，因为即使在实验室实验中也不能完全避免自然选择，而且对易于近交衰退性状的一些人工选择经常伴随着有意的近交。有两种选择效应必须区分开来：延缓接近固定和减少近交衰退。延缓接近固定在第 5 章末已讨论过。它导致纯合子的实际比例低于由谱系或群体大小计算的近交系数预测的比例。如果选择导致较好的等位基因在比偶然发生多的位点上固定，虽不延缓接近纯合性，近交衰退的降低也能发生。这样，为了有效利用选择，有必要同时近亲繁殖大量系群。于是选择可应用于系群之间，以便淘汰最坏的系群，保留最好的系群。其理由是遗传变异在系群内逐渐减少，在系群间逐渐增加。这将在下一章解释。

在抵消近交衰退中，选择的效果怎样呢？这是一个重要的实际问题，因为如果高度近交系要用于任何实践目的，则个体必须有适当的生存力和生殖力。在减少或甚至完全克服小鼠窝产仔数的近交衰退中，毫无疑问选择已是成功的。如果图 14-2（a）中的图形被外推至 F 为 100% 的完全近交，则高度近交小鼠预测的窝产仔数大约为 2。这是无选择时近

交的预期结果。实际上，大多数现存近交系的平均窝产仔数远大于2，表明选择在它们发育过程中成功地减少了近交衰退。Green（1968）列出的20个品系平均窝产仔数为5.7，范围从4.1到7.2。产生一个平均数高于原始群体的高度近交系的细节由下面的例子给出。

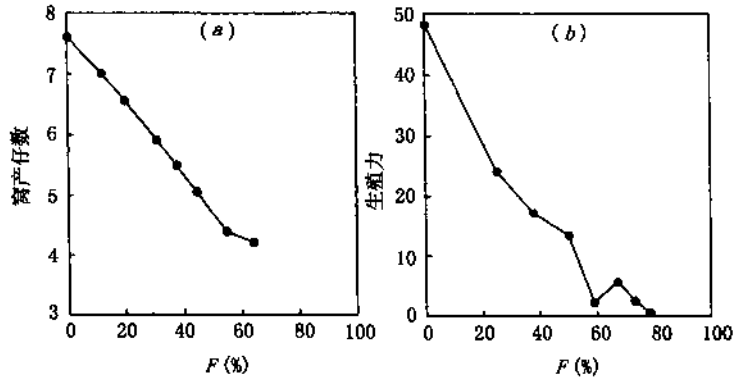


图 14-2 近交衰退影响生殖力的例子

（*a*）小鼠的窝产仔数（数据来自 Bowman 和 Falconer，1960）。第1窝出生时平均活仔数对仔鼠的近交系数作图。第一世代双表兄妹交配；此后全同胞交配。不进行任何选择。（*b*）果蝇（*Drosophila subobscura*）的生殖力。每天每对成年后裔的平均数对亲本的近交系数绘图。连续的全同胞交配（根据 Hollingsworth 和 Maynard Smith，1955）。

例 14.1

图 14-2（*a*）说明的窝产仔数的实验是以20个系群起始的。第一胎出生的所有雌性随后交配以提供平均窝产仔数。只饲养第一胎，因此，第一胎中没产出两种性别后代的系群就被淘汰。系群之间不应用任何其他选择。对10个系群应用了系群内选择（这些系群没有包括在图 14-2 中）。这种选择实际上对近交衰退没有影响。在第3世代（$F=55\%$），20个系群中有一个丢失了，在第6世代（$F=76\%$），只有3个系群继续存在。再经过5个世代，这3个系群平均窝产仔数为6.6，高出每窝2头的改进部分可归因于对20个系群中最好的3个系群的选择。这些系群中的一个系群无限期存活并成为一个叫做 JU 的“标准”近交系。在它经过20多世代近交之后再经过3年，第一胎中它的平均窝产仔数是9.0（McCarthy，1965），这要比近交前的原始群体的高。Bowman 和 Falconer（1960）和 Falconer（1960*a*）描述了这个实验。在小鼠的另外两个实验中也获得了相似的结果（Falconer，1971；Eklund 和 Bradford，1977）。

例 14.1 引用的3个实验证明，有可能得到小鼠的高度近交系，它的被选择性状至少与原始群体一样好。这提供了虽不是结论性但是很有力的证据，即超显性位点不是小鼠窝产仔数近交衰退的一个重要原因。如果一个超显性位点具有近交前群体的平衡基因频率，那么不可能有一个与非近交群体一样好的纯合系。但如果较好纯合子的频率低于非近交群体中的平衡值，那么在近交系中固定较好的纯合子可能增加平均值（Minvielle，1979）。有利于杂合子并影响近交衰退的选择方法不是易做的。这些方法取决于基因频率是否始于其平衡值，如果是这样，还取决于平衡频率是中间值还是极端值，即取决于两个纯合子适合

度大致相等还是有很大不同（Hill 和 Robertson，1968）。当起始频率是平衡值时，如果两个纯合子适合度几乎相等，那么选择通过延缓接近纯合性可减少近交衰退；但如果两个纯合子适合度很不相同，选择降低近交衰退的速率和总量，因为那时它引起较好的纯合子优先固定。

杂种优势

与近交衰退现象相补充的是其对立面，“杂种活力”或杂种优势。当近交系进行杂交时，后裔的那些以前因近交而下降的性状表现出增加。或者，采用一般的术语，近交时丢失的适合度在杂交中倾向于恢复。杂种优势的大小是杂交种和近交系平均值之间的差异。考察群体平均值如何取决于近交系数可以看出，简单地说，杂种优势就是近交衰退的反面现象，正如等式［14.4］显示的那样。如同先前所述，一个群体分成许多没有选择的近交系群，这样其平均基因频率是不改变的。如果系群随机杂交，杂种后代的平均近交系数可恢复到基础群体的大小，如果基因频率没有改变，则基因型的频率与基础群体的相同。因此，如果在系群之间随机进行许多杂交，那么可以预料到杂种后代中任何性状的平均值都与基础群体的群体平均值相同。换句话说，杂交的杂种优势预期与近交衰退相等。而且，如果在杂交之后群体因杂种和随后世代间的随机交配而得以继续，则近交系数将保持不变，因而预计群体平均值保持在基础群体的水平。这样根据理论推导，我们可以作出下列概括：在没有选择的情况下，在一个大群体中对系群进行近交，紧接着又进行杂交，预期是不会使群体平均值产生任何永久变化的。

例 14.2

小鼠的一个实验（Roberts，1960）被设计用来检验理论期望，即在没有选择的情况下，杂交的杂种优势应当等于近交时的衰退。研究的性状为窝产仔数。从一个随机交配群体中抽取 30 个系群，使它们通过 3 个连续世代的全同胞交配来繁殖，使近交系数在仔畜中达到 50%，在母本中达到 37.5%。近交过程中不施加任何选择，30 个系群中只有 2 个系群因为近交衰退的结果而丢失了。

表 14-3　小鼠不同阶段的平均窝产仔数

	近交系数		平均
	仔畜	母本	窝产仔数
近交前	0	0	8.1
近交	0.50	0.375	5.7
杂种仔畜	0	0.50	6.2
杂种仔畜和母本	0	0	8.5

近交三代后，系群间进行随机杂交，并且下一世代在 F_1 之间进行杂交以产生具有非近交仔畜的杂种亲本。在不同阶段观察到的平均窝产仔数给在表 14-3 中。近交衰退是 2.4，杂种优势是 2.8；两者在实验误差的限度内是相等的。

单 交

上面的理论结论指的是衍生于一个单一基础群体的系群之间大量杂交的平均情况。可是，实践中，经常感兴趣的却是稍有不同的问题，即两个系群之间特定杂交所表现的杂种优势，或没有已知共同来源的两个群体之间的杂种优势。在这些条件下将平均值的变化归因于近交系数的变化将是不合适的，而且按照两个系群的基因频率可以更好地表达杂种优势的理论基础。我们记得第 3 章指出近交导致系群之间基因频率的离散，随着近交的进行，系群的基因频率逐渐分化；近交系数是表示分化程度的一种方式（等式［3.14］）。从近交系数转到基因频率作为讨论的基础，因此我们要从杂交的一般或平均结果转到两个系群的特殊情况。

于是，让我们考察两个群体，称之为“亲本群体”，虽不必很大但两者都随机交配。亲本群体进行杂交以产生 F_1 或“杂种第一代”，F_1 个体在一起随机交配以产生 F_2 或“杂种第二代”。由 F_1 或 F_2 表现的杂种优势的大小将用与中亲值的离差，即与两个亲本群体平均值的差异来度量。首先考虑具有两个等位基因的单个位点的效应，基因频率在一个群体中为 p 和 q，在另一群体中为 p' 和 q'。让 y 代表两个群体之间基因频率的差异，因此 $y = p - p' = q' - q$。于是将第二个群体的基因频率 p' 和 q' 记为 $(p - y)$ 和 $(q + y)$ 可简化代数式。像前面一样，用 a、d、$-a$ 表示基因型值。假定在两个群体中基因型值相同，忽略上位互作。我们必须找出每一亲本群体的平均值和中亲值；然后找出 F_1 的均值和 F_2 的均值。亲本均值 M_{P_1} 和 M_{P_2} 可从等式［7.2］中获得。它们是：

$$M_{P_1} = a(p - q) + 2dpq$$

$$M_{P_2} = a(p - y - q - y) + 2d(p - y)(q + y)$$

$$= a(p - q - 2y) + 2d[pq + y(p - q) - y^2]$$

中亲值为：

$$M_{\bar{P}} = \frac{1}{2}(M_{P_1} + M_{P_2})$$

$$= a(p - q - y) + d[2pq + y(p - q) - y^2] \qquad [14.5]$$

当两个群体杂交以产生 F_1 时，从一个群体随机抽取的个体与从另一个群体随机抽取的个体进行交配。这等于随机地从两个群体中抽取基因，如表 14-4 所示。因此 F_1 的组成如表 14-5 所示。F_1 的平均基因型值是：

$$M_{F_1} = a(p^2 - py - q^2 - qy) + d[2pq + y(p - q)]$$

$$= a(p - q - y) + d[2pq + y(p - q)] \qquad [14.6]$$

用 F_1 和中亲值之间差异表示的杂种优势的大小，通过从等式［14.6］中减去等式［14.5］可获得：

$$H_{F_1} = M_{F_1} - M_{\bar{P}}$$

$$= dy^2 \qquad [14.7]$$

这样，就像近交衰退一样，杂种优势的发生取决于显性。无显性的位点（即位点的 $d = 0$）既不引起近交衰退也不引起杂种优势。两个特定系群或群体之间进行杂交的杂种优势大小取决于群体间基因频率差异（y）的平方。如果用于杂交的群体基因频率没有差异，就不会有杂种优势，当一个等位基因在一个群体中固定，另一等位基因在另一个群体中固定

时，则杂种优势将达到最大值。

表 14-4　F_1 的合子频率

			来自 P_1 的配子	
			A_1	A_2
			p	q
来自 P_2	A_1	$p-y$	$p(p-y)$	$q(p-y)$
的配子	A_2	$q+y$	$p(q+y)$	$q(q+y)$

表 14-5　F_1 的基因型

	基因型		
	A_1A_1	A_1A_2	A_2A_2
频率	$p(p-y)$	$2pq+y(p-q)$	$q(q+y)$
基因型值	a	d	$-a$

现在考察两个不同的亲本群体所有位点的联合效应。在基因型值归因于单独位点的加性联合的范围内，我们可以将由所有位点的联合效应产生的杂种优势表示为它们单独贡献的总和。这样 F_1 的杂种优势为：

$$H_{F_1} \approx \sum dy^2 \qquad [14.8]$$

从等式［14.8］可获得 3 个结论：

（1）如果一些位点在一个方向上是显性，而一些位点在另一个方向上是显性，那么它们的效应将倾向于互相抵消，可能观察不到杂种优势，尽管单个位点上存在显性。因此，像近交衰退一样，杂交时杂种优势的发生取决于定向显性，而缺乏杂种优势并不是单个位点无显性这个结论的充分根据。

（2）杂种优势的大小在某种程度上对每一个特定杂交都是特异的。两个系群因基因而不同，它们对所有系群配对都不会相同，因此不同的系群配对将有不同的 $\sum dy^2$ 值并将表现不同的杂种优势大小。

（3）如果杂交的系群是高度近交的，并且因而完全纯合，那么它们之间基因频率的差异只能是 0 或 1。于是等式［14.8］所显示的杂种优势是两个系群中有不同等位基因的那些位点的显性离差 d 的总和。

在我们接着考察 F_2 之前，或许值得注意，等式［14.7］和［14.8］中用基因频率差异的平方表示的杂种优势公式与以前用近交系数表示的近交衰退公式非常一致。如果我们考虑一个细分成很多系群的群体，并且我们假定各个系群随机配对，那么各对系群间基因频率的均差方将等于系群间基因频率方差的两倍，即 $\overline{y^2}=2\ \sigma_q^2$。基因频率的方差和近交系数之间的关系由等式［3.14］给出为 $\sigma_q^2=\bar{p}\,\bar{q}F$。因此 $\overline{y^2}=2\bar{p}\,\bar{q}F$，表明随机的系群对之间杂交的杂种优势平均大小等于等式［14.2］给出的近交衰退，但符号相反。

现在考察两个亲本群体的一个特定杂交的 F_2，此 F_2 是由 F_1 代个体之间随机交配产生的。随机交配的结果，F_2 的基因型频率将是相应于 F_1 基因频率的哈迪—温伯格频率。于

是应用等式［7.2］很容易得到 F_2 的平均基因型值。F_1 的基因频率，为两个亲本群体基因频率的平均值，一个等位基因是$\left(p-\frac{1}{2}y\right)$，另一个等位基因是$\left(q+\frac{1}{2}y\right)$。将这些基因频率分别代替等式［7.2］中的 p 和 q，可给出 F_2 的平均基因型值为：

$$M_{F_2}=a\left(p-\frac{1}{2}y-q-\frac{1}{2}y\right)+2d\left(p-\frac{1}{2}y\right)\left(q+\frac{1}{2}y\right)$$

$$=a(p-q-y)+d\left[2pq+y(p-q)-\frac{1}{2}y^2\right] \qquad [14.9]$$

F_2 表现的杂种优势的大小是 F_2 与中亲值之间的差异。因此，由等式［14.5］和［14.9］可得：

$$H_{F_2}=M_{F_2}-M_{\bar{P}}$$

$$=\frac{1}{2}dy^2$$

$$=\frac{1}{2}H_{F_1} \qquad [14.10]$$

因此我们发现 F_2 表现的杂种优势只有 F_1 表现的一半。换句话说，F_2 预期将 F_1 值向中亲值拉回一半。乍看这个结论似乎与早期得到的结论相矛盾，那时我们考察很多系群之间的杂交，于是 F_1 和 F_2 均值是相等的。两种情形之间的不同在于，由大量不同杂交间随机交配产生的 F_2 有和 F_1 一样的近交系数，而由单一杂交得到的 F_1 产生的 F_2 不可避免地会增加近交系数。如果用第 5 章描述的方式来计算近交系数，将会发现它是亲本系群近交系数的一半。因此从 F_1 到 F_2 的平均值变化可认为是近交衰退。这种衰退不能靠拥有大量的 F_2 亲本来克服，因为在只有两个系群的单交中或只有两个亲本群体的单交中已经形成了引起近交的群体大小限制。可是 F_3 和随后世代的近交系数不需要有进一步的增加。因此，假设没有其他引起基因频率变化的原因，那么随后世代的群体平均值将与 F_2 的相同。

当考察所有位点的联合效应时，假设不存在上位互作，那么 F_2 的预期杂种优势是 F_1 的一半同样是正确的。一单个位点的结论是建立在通过一单个世代随机交配即达到哈迪—温伯格平衡这一原理上。可是，这对同时考察多于一个位点上的基因型是不正确的，其理由已在第一章解释过。因此如果存在上位互作，群体平均值在 F_2 中将不会达到其平衡值，但是根据互作位点之间连锁的紧密程度，群体平均值在随后的世代中会或快或慢地接近其平衡值。上位互作的存在与度量尺度密切相关，但是这个问题直到第 17 章才讨论。这里我们只需注意到，由于与度量尺度有关，依理论根据预测的 F_2 杂种优势的减半，在实践中通常根本不会精确地发现，虽然 F_2 通常落在 F_1 和中亲值之间的某处图。14-3 列举出一些观察 F_1 和 F_2 世代杂种优势的植物实例。应当注意到 F_1 和 F_2 的某些性状值低于中亲值，因而杂种优势是负向的。这与我们定义杂种优势为 F_1 或 F_2 与中亲值之间的差异一点也不矛盾。差异的符号仅取决于度量的属性。例如，在图（e）到图（h）中表示的“门果成熟天数”这个性状表现了负向杂种优势；但如果性状称为“发育速度”并以时间的倒数表示，则杂种优势将是正向的。图 14-3 中不是所有的杂交都能提供对育种者有用的杂种优势。如果性状的迄纸镜是取值的唯一标准，那么杂交是无用的，除非 F_1 比两个亲本系群都好。术语杂种优势有时用来指有用的杂种优势，这是说，F_1 在数量上超过较好的亲本系群。

母体效应 在 F_1 和 F_2 世代中观察到的杂种优势的相对大小可能因母体效应变得复杂化。易受母体效应影响的性状，诸如产仔数，具有属于不同世代的两个组分。每一组分预

期遵从上面描述的 F_1 和 F_2 中杂种优势的一般模式，但这两个组分彼此之间步调不一致相差一个世代。这样 F_1 中观察到的杂种优势可归因于非母体部分，母体效应仍停留在近交水平上。可是，在 F_2 中，非母体部分将失去上面所说明的一半杂种优势，但因为母本现在处于 F_1 阶段，因此母体效应现在将表现其杂种优势的全部效应。这种相当复杂的情况可能从图 14-4 的图解中更容易领会。这指明了杂种优势如何出现在两个步骤中，第一步骤归因于非母体组分，第二步骤归因于母体组分。这两个步骤已在例 14.2 中阐明。

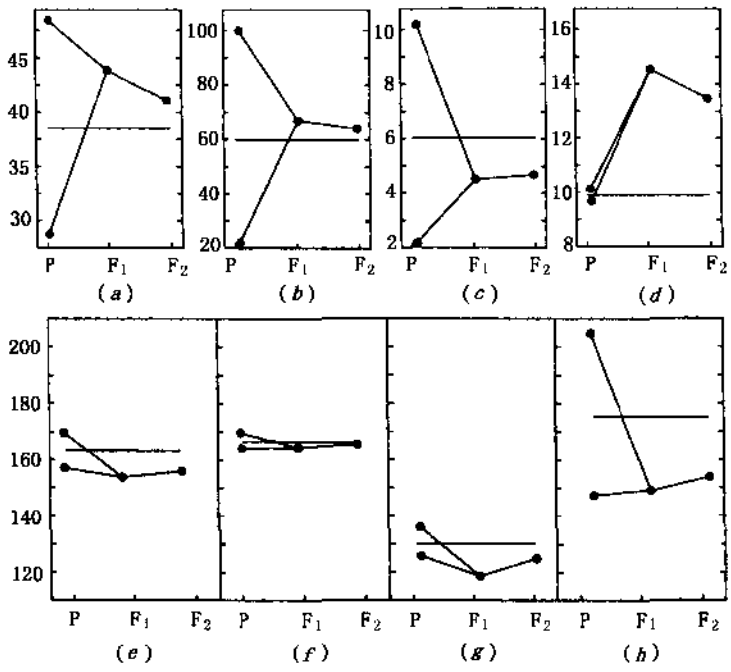


图 14-3　在植物高度近交系成对杂交中观察到的杂种优势的一些说明

点表示两个亲本品系、F_1 和 F_2 世代的平均值。水平线表示中亲值。图（*a*）指黄花烟草（*Nicotiana rustica*）（数据来自 Smith，1952）。其余的图指番茄（*Lycopersicon*）（数据来自 Powers，1952）。所代表的性状是：

（*a*）植株高度（英寸）；

（*b*）一个果实的平均重量（g）；

（*c*）每个果实的室数；

（*d*）每个室的平均重量（g）；

（*e*）～（*h*）4 种不同杂交中，播种至门果成熟之间以天表示的平均时间。

上位　我们已经看到一个特定杂交所表现的杂种优势的大小特别取决于杂交的两个群体之间基因频率的差异。这似乎表明杂种优势的大小随着两个群体间遗传分化程度的增加而增加，并且只受种间不育性的障碍所限制。可是，这是不真实的。通过适应当地条件而广泛分化的群体可能不表现杂种优势，而且可能在 F_2 世代中遭遇适合度的下降，正如在果蝇（*Drosophila*）群体研究中所看到的那样（Wallace 和 Vetukhiv，1955）。远缘杂交不表现预期的杂种优势可归因于上位互作，到目前为止我们一直假定上位互作不存在。对广泛不同的当地条件的适应包含许多不同的性状，因为一个生物体的适合度取决于其所有功能协调的相互关

系。因此很多位点上的基因是按它们对适合度的联合效应来选择的；以这种方式选择的基因的组合被认为是“相互适应”的。换句话说，一些适应性来自上位互作。当适应于不同条件的两个群体杂交时，杂交种对两种条件都不适应。而且，由于 F_2 代的分离，基因的有利上位联合丢失了。上位作用与杂种优势的关系由 Mayo（1987，p.148）作了汇总，并由 Geiger（1988）作了综述。图 14-5 和例 14.3 展示了玉米产量的杂种优势与杂交宽度有怎样的相关。

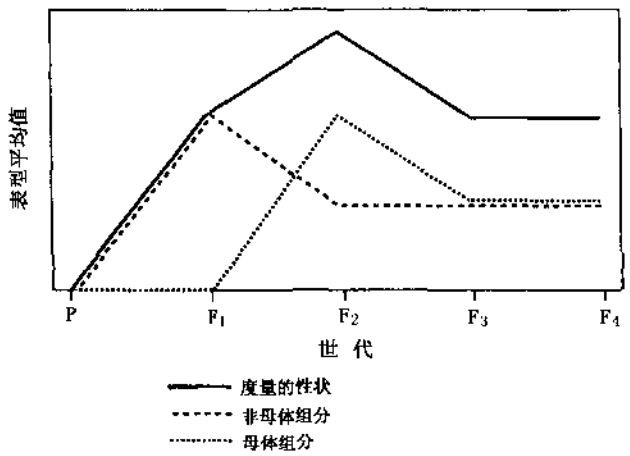


图 14-4 当两个系群进行杂交并且 F_2 和随后世代由随机交配产生时，在一个易受母体效应影响的性状中预测的杂种优势的图解。这里假定性状的母体组分和非母体组分分别表现相等大小的杂种优势，并当性状被度量时，通过简单地相加联合起来提供给性状

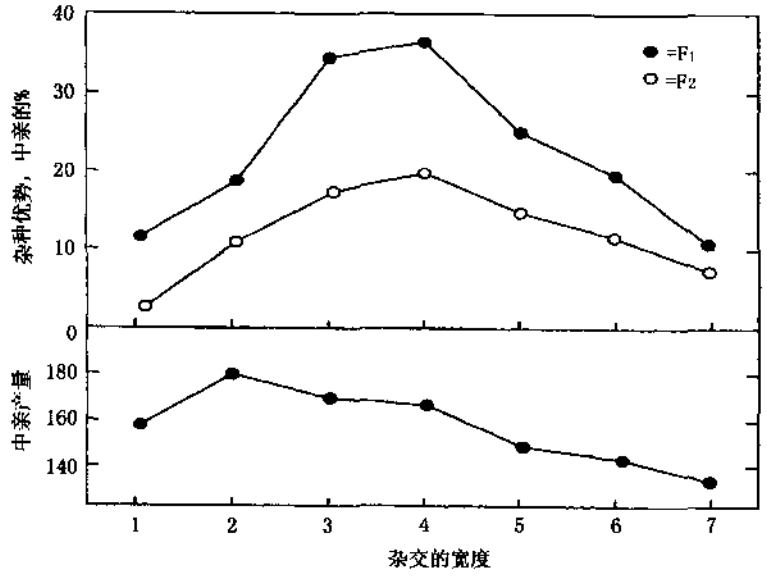


图 14-5 正如例 14.3 说明的那样，通过玉米产量指明的杂种优势和杂交宽度之间的关系

杂种优势以与中亲的百分率差异表示。下面的图表示以每株克为单位的中亲产量（数据来自 Moll 等，1965）。

例 14.3

（数据来自 Moll 等，1965）。玉米的品种或群体适应不同的地理区域。从 4 个区域的每一个内选择两个品种，进行了所有 28 种可能的杂交。每一个 F_1 个体之间随机交配产生 F_2 世代。在 4 个起源区域的每一个中种植品种和杂种。根据已知的品种祖先关系和区域的气候条件分为 7 个等级对杂种间多样性程度进行评估。等级 1 是同一区域的品种之间的杂种。图 14-5 显示了以亲本品种平均值的百分率表示的 F_1 和 F_2 代杂种优势对多样性程度绘成的图。也显示了亲本平均产量。中间“宽度”的杂种表现了最大的杂种优势。最宽的杂交给出的杂种优势低得多，尽管甚至在 F_2 世代中它仍是正向的。可是，从这些杂种中没有证据表明上位互作对互适性有贡献。如果上位互作对亲本群体产量有贡献，则 F_2 代杂种优势将不到 F_1 代杂种优势的一半；但在除等级 1 之外的所有杂种中，它多于一半（这些杂种的分析，见 Lynch，1991）。

习　题

14.1　假设玉米的一个随机交配群体具有每株 140g 的平均产量，且对产量具有已知效应的 4 个位点是分离的。如果基因效应由下表给出，那么一代自花授粉之后，由这些位点造成的近交衰退将是多大？每一种情形中基因效应是列出的基因型与“AA”纯合子之间在产量上的差异（g/株）；基因频率是等位基因“a”的。

位点	与 AA 的差异		q_a
	Aa	aa	
(1)	-10	-20	0.5
(2)	+5	-30	0.5
(3)	-20	-30	0.2
(4)	0	-60	0.1

[解答序号：69]

14.2　如果在习题 14.1 中指定的玉米群体中带选择的近交成功地固定了 4 个位点中每一个位点的较有利的等位基因，那么产量会增加多少？

[解答序号：79]

14.3　从例 14.2 有关小鼠窝产仔数的数据中计算由母本的和仔畜的近交引起的窝产仔数的近交衰退各是多少。假定与两者有关的 F 是线性的，并且母体和仔畜效应是加性联合的。如果小鼠能近交到 $F = 100\%$ 且没有任何选择操作，那么预计的窝产仔数将是多大？

[解答序号：89]

14.4　应用习题 14.3 的结果来预测当母本的近交系数是 0.56 以及仔畜的近交系数为 0.64 时窝产仔数的总近交衰退。这些是图 14-2（a）中描绘的实验最后世代的近交系数。这是一个与习题 14.3 中运用的例 14.2 不同的实验。这两个实验一致性程度如何？

[解答序号：99]

14.5 小鼠的一个对照系群保持了 30 个世代，尽管其有效群体大小不超过 40. 但其窝产仔数没有表现近交衰退的迹象。没有近交衰退的一个可能原因是，可能对窝产仔数有一些无意的选择。交配体系打算做到最低限度的近交，从每一家系中随机抽取一只青年雌鼠。理论上没有选择是可能的，但最低限度的近交不可能严格应用于实践中，因为一些家系包含没有存活的雌鼠。于是后备鼠不得不从其他家系中抽取，这些后备鼠可引起一些选择。需要多大的选择来抵消按习题 14.3 解答中预期的速率进行的近交衰退？后备鼠引起的选择将是个体选择，适合于个体选择的现实遗传力在同一品系中被估计为 0.22。

数据来自 Falconer，D.S.（1960）*J. Cell. Comp. Physiol.*，**56**（*Suppl.*1），153 ~ 167。

[解答序号：109]

14.6 在栽培番茄的品种之间进行杂交，正常情况下它们是自体受精的。由亲本品种和两种杂交的 F_1 在 3 周时期内产生的每株果实的平均重量（kg）如下：

杂交	P_1	P_2	F_1
(1)	1.44	1.36	1.41
(2)	1.28	0.88	1.42

F_2 世代以及由 F_2 自交产生的 F_3 世代的预测产量各是多少？

数据来自 Williams，W.（1960）*Genetics*，**45**，1457 ~ 1465。

[解答序号：119]

14.7 在习题 6.1 中计算了两个烟草品种杂交的 F_1 代和 F_2 代的平均叶片数。F_1 的平均值是 15.72 ± 0.24，F_2 的平均值是 15.84 ± 0.49。不知道亲本品种的叶片数，对 F_1 所表现的杂种优势你能得出什么结论？

[解答序号：129]

第 15 章　近交和杂交：Ⅱ．方差的改变

就数量性状的一般特性而言，从第 3 章对基因频率变化的描述中可清楚地看到近交对数量性状遗传方差的影响。而且，我们必须认为整个群体由许多系群组成。在近交或随机漂变的离散作用影响下，分离系群的基因频率趋向于极端值 0 或 1，并且系群在基因频率上变得有所分化。因为数量性状的平均基因型值取决于影响该性状位点上的基因频率，所以各系群在平均基因型值上变得有所分化或漂变。而且，因为方差的遗传组分随着基因频率趋向极端值而减小（见图 8-1），所以系群内的遗传方差也下降。因此，近交的一般结果是遗传方差的再分配；出现在系群均值之间的组分增加了，而出现在系群内的组分则下降了。换句话说，近交导致系群间遗传分化和系群内遗传一致。

一个近交群体细分为很多系群引入了一种附加的观测方差组分，即系群间组分，这对遗传方差组分的理论描述增加了相当的复杂性，这是不令人惊奇的。这里将只给出主要提纲的简要描述。详细的论述方法见 Wright（1969）以及 Weir 和 Cockerham（1977）。类似的是，系群杂交时，方差可分剖为杂种间和杂种内组分。杂种的方差将在本章末尾描述。遗传方差的再分配不是近交的唯一影响；实验已经证明环境方差有时也有影响。早在第 8 章，我们就提到过近交个体对变异的环境来源有较大的敏感性。这种现象干扰方差变化的实验研究，直到我们较好地理解了这一现象时，我们才对有关在观测的表型方差中清楚表示的遗传方差的理论期望值有充分的信任。与被考察的近交有关的另一个问题是高度近交系的遗传稳定性，这种稳定性对用于实验目的而涉及到的标准近交系的使用是重要的。

近　交

遗传方差的再分配

由加性基因（即无显性的基因）引起的方差再分配很容易推导。这是因为对加性基因来说，起始方差分配给系群内和系群间的比例并不取决于起始基因频率。可是，存在显性时，不了解起始基因频率，我们就不能推导方差的变化。这不仅增加了数学上的复杂性，而且致使一般的解决方法变得无能为力。我们首先考察加性基因的情形，然后非常简洁地表明对显性基因所获得的结论。选择的效果将不作特殊讨论。我们只需注意到自然选择将倾向于致使基因频率离散的实际状态不及依群体大小或谱系亲缘关系计算出来的近交系数所表明的那样高。

无显性　下面阐述的是指由加性基因引起的方差；它不适用于有显性的基因引起的加性方差。因此严格地说，结论只适用于没有非加性方差的性状。而且，它们可用来表明近交对方差的一般影响，也可作为对没有非加性遗传方差性状所预测的良好近似值。要给出的描述指的是缓慢近交，缓慢近交意味着两个连续世代中的近交系数没有太大的误差，可视为相同。可是，快速近交情况下的方差再分配除了开始的几个世代外没有很大的差异。

首先考察一单个位点。当无显性时，基础群体的基因型方差由等式［8.5］给出，为：

$$V_G = V_A = 2p_0q_0a^2$$

任一系群内方差是 $V_G = 2pqa^2$，其中 p 和 q 是该系群的基因频率。系群内平均方差是：

$$V_{GW} = 2(\overline{pq})\,a^2$$

其中（$\overline{pq}$）是所有系群 pq 的平均值。现在 $2(\overline{pq})$ 是整个群体中杂合体的总频率，由表 3-1 可知，它等于 $2p_0q_0(1-F)$，其中 F 是近交系数。因此：

$$\begin{aligned} V_{GW} &= 2p_0q_0a^2(1-F) \\ &= V_G(1-F) \end{aligned}$$

并且当对所有位点的方差求和时，这仍然成立。这样，系群内方差是（$1-F$）乘以起始方差，并且随着 F 接近 1，系群内方差接近于 0。

现在考察系群间方差。这是各系群真实均值的方差，可作为系群间组分从方差分析中估计。对一单个位点，仍无显性，对基因频率为 p 和 q 的一个系群的平均基因型值可从等式［7.2］获得，为 $M = a(p-q) = a(1-2q)$。这样我们要找出（$a-2qa$）的方差。假定上位作用不存在，因此 a 是固定的，即在所有系群中都相同。因此：

$$\sigma_M^2 = \sigma_{(2aq)}^2 = 4a^2\,\sigma_q^2 = 4a^2p_0q_0F$$

（来自等式［3.14］），它等于 $2FV_G$。当对所有位点求和时，这仍然成立。这样系群间遗传方差是 $2F$ 乘以基础群体的遗传方差。

表 15-1 当基础群体的遗传方差 V_G 是由无显性的基因引起时（近交系数为 F 的群体的方差分剖）

	(1)	(2)
系群间	$2FV_G$	$2fV_G$
系群内	$(1-F)\ V_G$	$(1+F-2f)\ V_G$
总和	$(1+F)\ V_G$	$(1+F)\ V_G$

注：f 是同一系群内个体的共祖率。(1) 缓慢近交，此时近似有 $F_{t+1} = F_t$，(2) 快速近交。

上面说明的遗传方差分剖的各个组分汇总在表 15-1 栏目（1）中。当近交快速时，各组分表示在栏目（2）中，Crow 和 Kimura（1970，p.138）解释了其理由。这里 f 是同一系群内个体的共祖率，如果个体在系群内随机交配，则它等于下一世代的近交系数。从表 15-1 我们看到整个群体的总遗传方差是系群内和系群间组分的和，它等于（$1+F$）乘以起始遗传方差。因此当近交完全时，整个群体的遗传方差会增加一倍，并全部表现为系群间组分。系群间的方差增加可由图 15-1 中实验数据来说明。

在近交完全之前，系群内遗传方差可在组成系群的家系内和家系间进行分剖。在系群内随机交配缓慢近交情况下，遗传方差可在全同胞家系内和全同胞家系间相等地分剖。系群内亲属的协方差就像第 9 章描述的，平均而言，每一系群都是一个独立的随机交配群体，具有总遗传方差 $(1-F)V_G$。从这一点，我们可以推导出任一系群内遗传力预期是多大。它是 $(1-F)V_G/[(1-F)V_G + V_E]$，这简化为：

$$h_t^2 = \frac{h_0^2(1-F_t)}{1-h_0^2F_t} \qquad [15.1]$$

其中，h_t^2 和 F_t 是在时间 t 时的系群内遗传力和近交系数，h_0^2 是基础群体的起始遗传力。这指明了随着小群体的近交预期遗传力如何减少。可是，公式仅适用于没有非加性方差的性状，且没有选择作用。自然选择的作用致使遗传力的下降低于预期值，特别是在缓慢近交情况下。这一点已为果蝇实验所证实（Tantawy 和 Reeve，1956）。

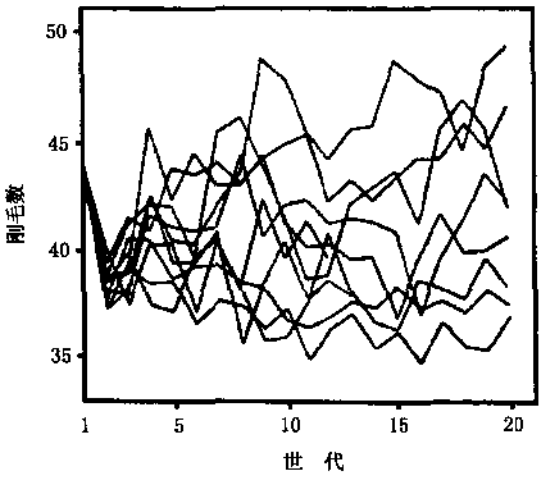


图 15-1　用黑腹果蝇（*Drosophila melanogaster*）腹部刚毛数显示的因随机漂变引起的系群间的分化

图显示出没有人工选择的全同胞近交中 10 个系群中每一个系群的平均刚毛数（根据 Rasmuson，1952）。

显性　我们将看到由加性基因引起的方差组分的变化不取决于基础群体的基因频率。可是，当我们考察具有任何程度显性的基因时，我们发现由近交引起的方差变化取决于起始基因频率，这使得用基础群体中存在的遗传方差给出通用的解答成为不可能。因此，当隐性等位基因处于低频率时，对于完全显性基因的情况，我们不能得出比 Robertson（1952）再多的结论。这最像是应用因有害隐性基因造成适合度变异的情况，尽管这里忽略了选择的效应。图 15-2 显示了由基础群体中频率为 $q=0.1$ 的隐性基因引起的方差再分配。图 15-2（a）指的是每一系群内只有一个家系的全同胞交配，图 15-2（b）指的是缓慢近交。一个令人惊奇的结论特征是系群内方差在近交开始时增加，当近交系数稍低于 0.5 时达到最大值，并以相当高的水平保持这一状况直到近交系数接近于 1。一般来说，近交开始阶段系群内变异增加这一明显反常的原因，可从图 8-1（b）显示的基因频率和由显性基因引起的方差之间的关系中看出。视基因频率开始值为 0.1，随着近交的发生，它将在一些系群中增加，在其他系群中减少，平均而言增加的数量等于减少的数量。但检测图示看出，与降低同样数量的基因频率而减少的方差相比，数量相同的基因频率增加造成的方差增加要更多一些。因此，平均而言，系群内方差在近交的早期阶段将增加。只有当遗传方差的相当可观部分是由低频率的隐性基因引起时，方差的这种增加在实践中才可查明。

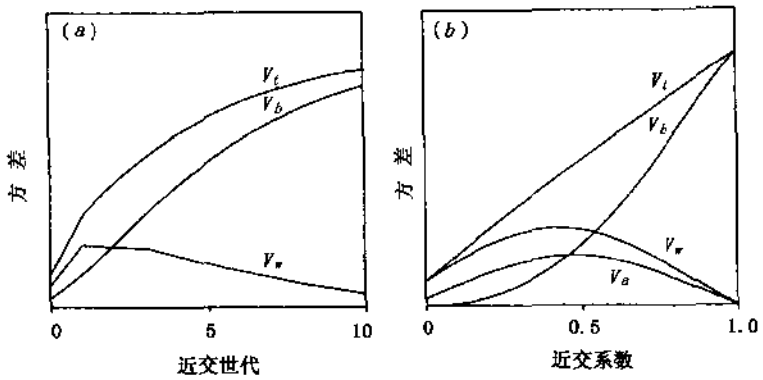


图 15-2 由起始频率 $q_0=0.1$ 的一单个完全隐性基因引起的方差再分配

（a）全同胞交配，（b）缓慢近交（根据 Robertson，1952）

V_t = 总遗传方差；

V_b = 系群间组分；

V_w = 系群内组分；

V_a = 系群内加性遗传方差。

环境方差

前几章我们多次提到对不同基因型方差的环境组分可能不同的事实；特别是近交个体经常比非近交个体表现出更多的环境方差。这个事实已为比较近交种和杂交种的很多实验所揭示。如果高度近交的系群比它们之间的 F_1 杂种（即杂交种）更易于变化，那么这种差异一定归因于方差的环境组分，因为杂交种和近交系群的遗传方差在数量上是可以忽略不计的。近交种比杂交种对变异的环境来源有较大的敏感性不是一个普遍现象，但这种现象可在大量的性状和生物体中观察到。一些实例引用在表 15-2 中；Lerner（1954）和 Wright（1977）给出了另外的例子。在大鼠和小鼠的行为性状中已广泛研究了这种现象；Hyde（1973）综述了这些实验研究，他发现 19 种行为性状中有 14 种表现了这种现象。

在我们考察环境方差不同的可能原因之前，应当提及与这种现象有关的两点。第一点是技术问题。如果近交种和杂交种之间的性状平均值不同，并且经常是这样，于是很难确定一个合适的基础来进行方差的比较。因而有必要确信，所发现的方差任何差异不仅仅是平均值差异的一种反映。换句话说，必须找到一种与平均值不相关的变异的度量。变异系数经常是一种适宜的度量，虽然不总是。这个问题基本上是尺度选择的问题，将在第 17 章再考虑此问题。第二点涉及被度量的环境变异的性质。要区分由环境引起的变异和生物体对特定环境条件的适应性反应。当考察变异与适合度的可能关系时，这种区分是必要的。在第一种情形下，假定存在一种最适基因型，它对考察的所有环境条件都是相同的。哺乳动物的体温就是一个明显的例子。于是对环境变量的感觉迟钝性是适合度的一个方面：最适应的个体是那些能调节它们的发育或生理过程，以便实现最适表现型的个体，而不管环境变量低于最适值。用这种方式对变异的限制称为*体内平衡*。引起变异的环境因素不一定是外部的，而且也包括发育变异的内在原因。在第二种情形下，当生物体有适应反应时，在不同的环境中性状有不同的最佳状态。与环境温度变异有关的排汗就是一个例

子。对这类性状，与适合度的关系是颠倒的：具有变化能力的个体比没有这种能力的那些个体更能适应。

表 15-2　近交种环境方差大于杂交种的性状实例

		近交个体	杂交种
黑腹果蝇(*Drosophila melanogaster*)(Robertson 和 Reeve,1952)			
翅长:6 个近交,6 个 F_1		2.35	1.24
小鼠(McLaren 和 Michie,1956)			
戊巴比妥钠麻醉的持续期(log 分钟)			
2 个近交,1 个 F_1		0.0665	0.0165
小鼠(Yoon,1955)			
阴道开张日龄(天)			
3 个近交,2 个 F_1		51.7	17.4
小鼠(Chai,1957)			
给定日龄的体重(g)(C.V.)^2	出生	119	59
2 个近交,1 个 F_1	21 天	98	47
	60 天	24	19
大鼠(Livesay,1930)			
90 日龄体重(C.V.)^2			
3 个近交,2 个 F_1		522	170
玉米(Shank 和 Adams,1960)			
株高(C.V.)^2		44	30
穗重(C.V.)^2		412	198
10 个近交,5 个 F_1			

注:数值是对近交系和 F_1 平均的表型方差。F_1 均为多次杂交的。(C.V.)^2 = 变异系数的平方。

那么一些性状在近交种中比杂交种中更易于变化的原因是什么呢？这种现象可看作是近交衰退或杂种优势的一种表现（Mather，1953）。基于这一点，表现这种现象的性状其体内平衡是有利的，近交已减少了适合度的这种组分，当杂交种与近交种比较时，这种组分引起近交种中环境方差增加并相应地减少杂交种的环境方差。这种解释暗示存在影响变异性，即体内平衡的基因，这些基因独立地对性状均值有任何影响，而且这些基因一定有某种程度的增加体内平衡的定向显性。用生化术语可对这种现象做另一种解释。酶的不同等位形式处于诸如温度或 pH 等环境变量不同的数值上，经常有其最大的酶活性。因此具有两种酶形式的杂合体，可能比仅有一种酶形式的纯合体，在一个较宽范围的环境变异中保持充足水平的酶活性。就反映在被度量性状上的酶活性而言，杂合体不及纯合体对环境变量敏感。基于这一点，稳定性并不一定与适合度相关，而它仅仅是影响性状平均值的某些位点上杂合性的生化结果。因此就这些位点对变异性的影响而言，它们是超显性的，尽管它们的影响不必定在性状平均值上或是在适合度上。

近交系的一致性

实验动物，特别是小鼠的近交系作为实验材料广泛应用于检验和分析以及生物学研究的许多其他领域。近交系是单一系群，因为想获得一致性而使用。对一些用途，例如当不需要抗原差异时，需要的正是遗传一致性。丰富的经验已表明实验小鼠的高度近交系完全

满足遗传一致性的这种要求。可是，对另一些用途，需要的不仅仅是遗传一致性，还有表型一致性。动物变异越小，在度量它们对处理的反应过程中为达到一定精确度而需要使用的动物数就越小。因此在这个方面近交动物的数值取决于最初性状有多少方差是遗传的，因为只有遗传方差才可通过近交消除。它也取决于近交影响的环境方差有多大；正如表15-2例示的，增加的环境方差有时会抵消减少的遗传方差以致于近交系在表型上比非近交家畜更易于变异。当然，获得遗传一致性又不增加环境变异的方法是使用两个近交系间杂交的F_1。这具有额外的好处，F_1代个体通常比近交种健康、易于生存和更易于生育，尽管它不降低生产费用，因为必须保持近交种作为亲本。也必须提及与近交种或F_1代动物使用有关的另一点。一个近交系或两个近交系的F_1代有一个唯一的基因型；而且，一个近交系具有唯一基因型这种情况在自然群体中是不可能发生的。因此检验对一个近交系或一个杂种系的任何处理的反应就是检验对一个基因型的反应。基因型之间可能存在可察觉的反应差异，因此从一个系或杂交中获得的结果用于其他系便不一定可信了。

突 变

在第12章我们看到，突变产生了数量性状的新变异，并且这种变异在长期选择的背景中是不能忽略的。突变对有关近交后果所得结论的影响有多大？首先考察单个位点上的选择中性的突变体。等位差异的起源速率在第14章已指出，是简单地等于中性等位基因的突变率，而不考虑群体大小。在现在的背景中重复这种讨论：在同胞交配的一个系群中每一常染色体位点有4个代表，因此在一特定位点上突变发生率是$4u$，其中u是每世代每一配子的突变率。但已发生的突变有$\frac{1}{4}$的机会被固定。因此每世代等位替代率等于突变率。不严格的中性突变体会变为固定，尤其在很小的群体中。在第4章陈述了，如果对突变体的不利选择系数低于$\frac{1}{2N_e}$，在同胞交配系群中该选择系数大约是0.2（见第5章），那么该突变体是“有效中性”。因此，例如，如果不利选择系数低于大约20%的等位基因的突变率为10^{-5}，并且如果位点总数是10 000，于是平均而言，在任一亚系群中每10个世代将有一次等位替代（$10^{-5}\times 10^{4}\times t=1$；$t=10$）。对于同胞交配，一个新突变体很快被固定或丢失，因此在任一时刻分离的位点数是很小的。这样，关于同胞交配系群的遗传一致性得出的结论极少受突变的影响。可是，当我们考察数量性状时，近交系的长期不变不是绝对的，突变是不能忽略的，正像现在我们所看到的那样。

不知道影响性状的位点数目，也不知道位点效应的大小，我们不能将前面的讨论延伸到数量性状上，我们对这两种情况知之不多。可是，重要的事情是所有相关位点上由突变产生新方差的速率，以及这种速率的一些估计值是可获得的。一个世代中由突变产生的方差可称为*突变方差*，用符号V_M表示。作为环境方差V_E的一个比例来表达是最方便的。若干物种中各种性状的估计值表明，平均而言V_M具有$V_E\times 10^{-3}$数量级（这些估计值由Lynch，1988综述，并在第20章再次考虑之）。每一世代由突变引起的这种新变异累积起来，直到每一世代的获得量被近交固定造成的损失所平衡。此后遗传方差保持不变。对极度近交，这种恒定的状态仅在几个世代之后就可达到（见Lynch和Hill，1986）。

每一世代损失的遗传方差数量可推导如下。系群内保持的起始方差的比例由表15-1给出为（$1-F$）。如果我们借助等式［3.12］用ΔF的形式来表示F，并借助等式［4.1］

用 N_e 的形式来表示 ΔF，同时如果我们让 $t=1$，我们就获得 $1-\left(\frac{1}{2N_e}\right)$ 作为一个世代后保持的比例。如果我们用 V_G 表示任一世代存在的方差，那么下一世代损失的数量是 $V_G/2N_e$。当这与新的突变方差 V_M 相等时，我们发现：

$$V_G = 2N_eV_M$$

这是一个近交系群内恒定的遗传方差数量。它指的是对性状只有加性效应的中性基因，虽然非加性几乎不产生变异（见 Lynch 和 Hill，1986）。对同胞交配，$N_e=2.6$（见第 5 章），并且如果我们代入这个值和上面给出的 V_M 的估计值，我们发现 $V_G=(5\times10^{-3})V_E$（近似）。如果用遗传力表示，这可能更有意义，即是 $V_G/(V_G+V_E)=0.5\%$。因此，突变不影响这个结论：对最实用的目的而言，同胞交配系群在遗传上是一致的。

亚系歧化

与突变对同胞交配系群内方差的影响很小形成对比，如果一个近交系分裂成多个亚系，则突变确实有重要的影响。标准近交系的使用方便了来自不同实验室结果的比较。离散到不同实验室不可避免地导致系群分裂成亚系，知道因遗传漂变对突变方差的作用造成亚系分化的程度是重要的。

如果系群在分离成亚系之前是高度近交的，则基础群体的起始方差一点也留不下，而且，源于突变的方差将达到其稳定水平 $2N_eV_M$，这也将是它们分离后亚系内的稳定水平。每一世代中亚系间方差增加的数量如下。亚系间方差是 $2FV_G$（根据表 15-1）。用 N_e 的形式来表示 F，就像上面对系群内方差所进行的变换那样，这给出每一世代亚系间方差增加的速率为 V_G/N_e。代入 $V_G=2N_eV_M$，我们发现系群间方差每世代增加 $2V_M$。（这种速率独立于群体大小，因为较大的群体有很大的系群内方差，而且由随机漂变引起的分化更慢。）如果亚系是在 t 世代以前分离的，那么亚系均值的方差将是 $2tV_M$，因为 $2V_M$ 是每一世代的分化速率。当然，这假定亚系之间没有环境差异。为了正确判断这种分化的实际重要性，我们可以将亚系间方差 σ_b^2 表示为表型方差 σ_W^2 的一个比例。于是有：

$$\frac{\sigma_b^2}{\sigma_W^2}=\frac{2tV_M}{2N_eV_M+V_E}$$

如果我们代入 $V_M=10^{-3}V_E$ 和 $N_e=2.6$，得到：

$$\sigma_b^2/\sigma_W^2=(1.63\times10^{-3})t$$

因此，例如 25 个世代以前分离的同胞交配亚系均值之间的方差将大约是亚系内表型方差的 5%。因此，亚系之间可查明的差异是可预测的，即使差异很大。实际上已在几个实验中发现了它们，例如 McLaren 和 Michie（1954）、Grewal（1962）、Festing（1973）以及下面例 15.1 描述的实验。有关突变和亚系歧化的进一步细节见 Lynch 和 Hill（1986）以及 Lynch（1988）。

例 15.1

这是 Bailey（1959）进行的小鼠两个近交品系的一项研究的简要描述。在 C57BL/6 品

系的两个亚系和 BALB/cAn 品系的两个亚系中研究了各种骨骼大小的 6 种度量。这两个品系分别通过全同胞交配 30 个世代和 78 个世代之后分离成亚系，因此由残余的分离引起的分化可能不是很大。从亚系的分离以来，世代的平均数是，C57BL 为 46.5，BALB 为 9。于是两个品系的亚系在一些性状方面歧化很严重，两个品系的亚系在 6 个性状中有 4 个性状差异显著。C57BL 中 6 个性状的 σ_b^2/σ_W^2 比率从 0 变化到 1.27，平均值为 0.41，而 BALB 中该比率从 0.08 变化到 0.52，平均值为 0.23。通过对上面给出的表达式 σ_b^2/σ_W^2 代入适当的值并重新整理，则可获得突变方差的估计值。C57BL 中 $(V_M/V_E \times 10^3)$ 的值从 0 变化到 15，平均值为 5，而 BALB 中从 4 变化到 34，平均值为 14。可是，这些估值是很不准确的，因为它们仅以每一品系的两个亚系为基础。

杂 交

在本章的第一部分，我们看到一个数量性状的遗传方差是如何在近交系间和近交系内分配的。现在我们必须考察方差是如何在杂种间和杂种内分配的这个补充的问题。杂种之间的方差对预测有多少改良预计是来自近交和杂交是重要的；这将在下一章解释。由于下列原因，与近交之后的方差相比，杂交之后的方差是一个较为简单的问题。只要选择没有改变基因频率，则近交系产生的配子就不同于非近交群体产生的配子。这意味着任一 F_1 个体的基因型原则上都可在一随机交配群体中找到；反之，起始随机交配群体中的任何基因型也能在杂种之间找到。因此，杂交之后的总遗传方差与基础群体中的相同。同一杂交的 F_1 个体可认为是一个“家系”，该家系具有的亲缘程度取决于亲本系群的近交系数。这些“家系”的协方差就是杂种之间的方差。如果亲本系群是完全近交的，则同一 F_1 的所有成员有同一的基因型，并且杂种之间的方差等于基础群体的总遗传方差。如果亲本系群不是完全近交的，那么从共祖率获得的亲缘系数中可以求出家系的协方差。现在将通过随机杂交来解释这一点。

杂种之间的方差

假定相当多的系群是近交的，且没有选择来自同一基础群体，所有系群具有相同的近交系数。然后在系群之间随机进行杂交。进一步假定每一个杂交由亲本系群的很多个体组成，这些亲本个体与它们系群的每一个其他个体不相关。这意味着系群内的遗传方差完全表现在杂种中。于是，每一个杂种可认为是一个家系，其成员之间具有一定程度的亲缘关系。杂种之间的方差组分就是这些相关个体的协方差。亲属协方差的组分以遗传组分以及亲缘系数 r 和 u 的形式在等式［9.13］中给出。系数 r 关联加性方差，u 则关联显性方差。系数 r 和 u 可从等式［9.11］和［9.12］显示的共祖率中获得。在一个杂种中有关的共祖率如下。有关个体显示在图 15-3 中，其中表示它们的字母与图 15-4 中的那些相符合。A 和 C 是一个亲本系群的两个个体；B 和 D 是另一个亲本系群的两个个体。P 和 Q 分别是由 A 和 B、C 和 D 杂交产生的两个 F_1 个体。有很多其他的杂交可产生具有像 P 和 Q 谱系亲缘关系的其他 F_1 个体。我们需要适合于 P 和 Q 这种亲缘关系的用共祖率 f_{PQ} 表示的协方差。首先，根据等式［9.11］，$r = 2f_{PQ}$，根据等式［5.2］，$f_{PQ} = \frac{1}{4}(f_{AC} + f_{AD} + f_{BC} + f_{BD})$。这两个亲本系群无亲缘，因此 $f_{AD} = f_{BC} = 0$。这两个系群有相同

的近交系数，因此 $f_{AC} = f_{BD} = F$，其中，F 是系群下一世代的近交系数，即与 F_1 世代同期的世代。这样 $f_{PQ} = \frac{1}{2}F$，这给出：

$$r = F \qquad [15.2]$$

下一步，根据等式 [9.2]，$u = f_{AC}f_{BD} + f_{AD}f_{BC}$，它可化简为：

$$u = F^2 \qquad [15.3]$$

通过将等式 [15.2] 和 [15.3] 代入等式 [9.13] 中，现在我们可用近交系数的形式来表示协方差。这给出杂种之间的方差为：

$$\sigma_X^2 = FV_A + F^2V_D + F^2V_{AA} + F^3V_{AD} + F^4V_{DD} + \cdots \qquad [15.4]$$

在这个表达式中，V_A 和 V_D 是基础群体的加性和显性方差；V_{AA}、V_{AD} 和 V_{DD} 是第 8 章所解释的互作组分；F 是近交系数，不是杂交中所用亲本的，而是系群下一世代的近交系数。遗传方差的剩余部分表现在系群内，即 $(1-F)V_A + (1-F^2)V_D$ 等。杂种之间的方差是真实平均值的方差，它可以作为方差组分从方差分析中估计。等式 [15.4] 给出了这个组分的遗传容量。要是杂种之间没有环境差异，它就等于实验上估计出来的大小。如果要计算杂种的观察均值的方差，那么这将增加杂种内方差的一个部分，这取决于每一杂交中度量的个体数，其理由在第 13 章中有关家系选择的部分做了说明（见表 13-3）。

等式 [15.4] 中值得注意的要点是遗传方差不同组分的贡献在不同程度上与近交系数相关联。加性方差的贡献随着 F 的增加而线性增加，但显性和互作组分的贡献以 F 的平方或较高次幂的比例而增加。结果是如果性状具有大量的非加性方差，随着 F 接近 1，杂种在近交的后期要比在早期分化得快得多。当 $F = 0.5$ 时，每一杂种在遗传上等于基础群体的一个全同胞家系；当 $F = 1$ 时，整个的遗传方差出现在杂种之间。杂种间方差与 F 相关联方式的实践重要性将在下一章考虑。

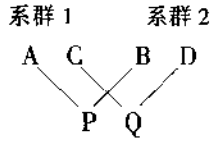


图 15-3

配合力

上面获得的杂种之间的方差是以大量系群随机杂交为假设的。“大量”意味着每一个系群只在一个杂交中使用。与此相反，如果每一系群和几个其他系群杂交，那么杂种之间的方差就能分剖，这一点对理解杂交育种在改良中的用途有很大的重要性。为便于解释，我们将假定使用大量的系群、杂种和个体，以便所有平均值可准确无误地估计。一个系群与几个其他系群进行杂交提供了该系群的一个附加的度量，即该系群在它的所有杂种中的平均表现。这个平均表现，当以与所有杂种平均值的离差表示时，就称为该系群的一般配合力。一般配合力是作为一个亲本的这个系群所有 F_1 个体的平均值，该数值以与杂种的全群平均值的离差形式表示。于是，任何特定杂种都有一个期望值，它是两个亲本系群一般配合力的和。可是，杂种可能在较高或较低程度上偏离这个期望值。这个离差就称为这两个系群配合的特殊配合力。采用统计术语，一般配合力是主要效应，特殊配合力是互作效应。因此，系群 P 和 Q 之间一次杂交的真实均值 X 可表示为：

$$X - \overline{X} = GCA_P + GCA_Q + SCA_{PQ} \qquad [15.5]$$

其中，$\overline{X}$ 是所有杂种的平均值，GCA 和 SCA 分别是一般配合力和特殊配合力。实践中，

另一项 E 必须加在右边以表示估计 X 的抽样误差。

等式［15.5］右边的各项互不相关，因此杂种之间总方差（不包括误差方差）组成如下：

$$\sigma_X^2 = \sigma_{GCA}^2(\mathrm{M}) + \sigma_{GCA}^2(\mathrm{F}) + \sigma_{SCA}^2 \qquad [15.6]$$

其中，（M）和（F）分别指的是雄性亲本和雌性亲本系群的一般配合力。如果系群在性别和任何其他方面都没有区别，那么等式［15.6］变为：

$$\sigma_X^2 = 2\,\sigma_{GCA}^2 + \sigma_{SCA}^2 \qquad [15.7]$$

杂种间总方差分剖成的两个组分是一般配合力方差和特殊配合力方差。根据第 9 章说明的意思，这些是方差的观察组分，可从方差分析中估计。它们的重要性就在于这个事实：遗传方差的有因组分对它们的贡献不同，正象我们现在看到的那样。

有一个系群作为共同雄性亲本的一组杂种可认为是与父系半同胞家系类似的家系。这些家系的协方差就是雄性系群一般配合力的方差 $\sigma_{GCA}^2(\mathrm{M})$。可以用同先前一样的方法从亲缘系数中获得这个协方差。图 15-3 用来说明相关系谱。个体 A 和 C 是共同雄性系群的两个成员，但现在 B 和 D 是与 A 和 C 进行杂交的两个不同系群的成员，产生两个 F_1 个体 P 和 Q。现在除 $f_{AC} = F$ 外，所有共祖率都为 0。这给出 $r = \frac{1}{2}F$ 和 $u = 0$。代入以前的等式［9.13］有：

$$\sigma_{GCA}^2(\mathrm{M}) = \frac{1}{2}FV_A + \frac{1}{4}F^2V_{AA} + \cdots$$

相同的论述适用于作为雌性亲本的系群，因此，如果不存在母体或性连锁效应，则 $\sigma_{GCA}^2(\mathrm{F}) = \sigma_{GCA}^2(\mathrm{M})$。特殊配合力的方差是等式［15.4］给出的杂种间总方差的剩余部分。因此我们获得下面的杂种方差组分的组成：

一般配合力：

$$\left.\begin{aligned}
&\text{雄性亲本：}\quad \sigma_{GCA}^2(\mathrm{M}) = \frac{1}{2}FV_A + \frac{1}{4}F^2V_{AA} + \cdots \\
&\text{雌性亲本：}\quad \sigma_{GCA}^2(\mathrm{F}) = \frac{1}{2}FV_A + \frac{1}{4}F^2V_{AA} + \cdots \\
&\text{特殊配合力：}\quad \sigma_{SCA}^2 = \frac{1}{2}F^2V_{AA} + F^2V_D + F^3V_{AD} + F^4V_{DD} + \cdots
\end{aligned}\right\} \qquad [15.8]$$

从这里可以看出，一般配合力的差异是由基础群体中加性方差和 A × A 互作引起的；而特殊配合力的差异则归因于非加性遗传方差。因此，一般配合力方差随着 F 而线性增加（互作组分除外），而特殊配合力方差随着 F 的较高次幂增加。因此，可以预料，随着近交达到高水平，增加更快的是特殊配合力方差，而不是一般配合力方差。

等式［15.8］中的遗传方差组分是这样一种随机交配群体的遗传方差组分，它的所有基因频率都与杂交系群的相同且相引和相斥连锁处于平衡状态。这种随机交配群体可认为是真实或假想的基础群体，从中可以无选择地衍生出系群。或者，换个说法，这种随机交配群体也可认为是由杂种间随机交配产生的一个合成群体，然后通过随机交配繁殖足够长时间以达到连锁平衡。要被采用的这些观点中的哪一个都会影响方差分析的细节。在第一种情形中，系群被认为是群体的一个样本，因而是随机因素；在第二种情形中，系群是整

个群体，因而是固定因子（详细细节见 Griffing，1956*b*）。

配合力的估计 估计一般配合力的方法在植物中运用是方便的，它被称为多系天然杂交（多向杂交）方法。将要测验的来自所有系群的大量植株种植在一起，或用异花授粉的自然机制或用小区内植株的排列方式阻止自花授粉，令其自然授粉。因此来自一个系群植株的种子是与其他系群随机杂交的混合物，当种植它们时它们的表现测定了该系群的一般配合力。测定的一般配合力是作为雌性亲本系群的一般配合力。如果假定雄性和雌性亲本一般配合力的方差相同，那么一般和特殊配合力的方差可根据等式［15.8］进行估计和解释。

一个系群的一般配合力也可通过它与基础群体的个体而不是其他近交系的个体进行杂交来估计，这种方法被称为顶交。正如先前注意到的那样，它等于与来自没有选择的基础群体的一组随机的近交系进行杂交，因为来自近交系的配子在遗传含量上与基础群体中的配子没有差异。

对近交系杂交，一个普遍使用的试验设计是双列杂交，其中每一个系群都与每一个其他系群进行杂交。下面例 15.2 解释和例证了对各个系群一般配合力的估计。为估计方差目的而进行的双列杂交分析是复杂的，因为它取决于是否包括反交以及对所指群体方差的遗传组成作出的假设。理论和分析方法由 Griffing（1956*a*，*b*）所解释并由 Pooni、Jinks 和 Singh（1984）以及 A.J.Wright（1985）所评估。

例 15.2

采用 Sprague 和 Tatum（1942）关于玉米近交系之间杂交的数据来说明一般和特殊配合力的计算。使用 10 个系群，每一系群都与其他 9 个系群杂交，但没有进行反交。因此有 $\frac{1}{2}n(n-1)=45$ 个杂交组合，n 是系群数。产量给在表 15-3 中，这些是每个杂交组合的平均产量，以每英亩多少蒲式耳为单位（每英亩 100 蒲式耳玉米 = 每公顷 6.725 吨）。标题为 T 的栏目给出了每一系群在所有 9 个杂交中的总产量，这可通过对合适的栏并沿着箭头所指示的行进行求和而获得。注意到每一个杂交组合贡献两个总数，因此 $\sum T=2\sum X$，其中 X 是一个杂交组合的产量。

如果有极其大量的系群，那么一个系群的一般配合力可简单地计算为它的均值即 $T/(n-1)$ 与全群平均值 $\overline{X}$ 的离差。可是，对小数目的系群，这不是有效的，因为每一个其他的系群贡献它的一般配合力的一个分数即 $1/(n-1)$ 给被测系群的平均值。这样，例如，系群 A 在所有它的杂交组合中的平均值 $\overline{A}$ 是：

$$\overline{A}-\overline{X}=G_A+\frac{1}{n-1}(G_B+G_C+\cdots+G_N)$$

其中，G 是由下标指明的系群 A 到 N 的一般配合力。现在 $\sum G=0$，因此 $(G_B+G_C+\cdots+G_N)=-G_A$，因而：

$$\overline{A}-\overline{X}=\left(\frac{n-2}{n-1}\right)G_A$$

从中可得：

$$G_A=\left(\frac{n-1}{n-2}\right)(\overline{A}-\overline{X})$$

用总数比用平均值计算更方便。代入 $\overline{A} = T_A/(n-1)$和 $\overline{X} = \sum T/n(n-1)$得：

$$G_A = \frac{T_A}{n-2} - \frac{\sum T}{n(n-2)}$$

表中系群的一般配合力列入标题为 *GCA* 的栏中。Simmonds（1979，p.112）给出了适合于其他双列杂交设计的公式。

表 15-3 玉米杂交种的产量

	B	C	D	E	F	G	H	I	J	*T*	*GCA*
A	86	84	98	98	92	92	97	81	88	816	3.75
B	→	91	105	102	86	92	79	80	90	811	3.125
C	·	→	87	80	65	84	93	77	83	744	−5.25
D	·	·	→	97	100	101	97	91	80	856	8.75
E	·	·	·	→	97	83	93	78	83	811	3.125
F	·	·	·	·	→	80	93	76	70	759	−3.375
G	·	·	·	·	·	→	90	74	72	768	−2.25
H	·	·	·	·	·	·	→	91	96	829	5.375
I	·	·	·	·	·	·	·	→	78	726	−7.50
J	·	·	·	·	·	·	·	·	→	740	−5.75
总和							$2\sum X = \sum T = 7\,860$				0.000
全群平均值							$\overline{X} = 87.333$				

现在能计算每一个杂交组合的期望值，“期望”意味着在一点也不了解特殊配合力的情况下，数值可从两个一般配合力来预测。以最好的杂交组合 BD 为例，$E(X_{BD}) = 3.125 + 8.75 + 87.333 = 99.21$。观察值和期望值之间的差异估计了两个系群配合时的特殊配合力：$SCA_{BD} = 105 - 99.21 = +5.79$。这样获得的特殊配合力的数值服从于估计 X_{BD}的抽样误差。图 15-4 展示了观察产量与期望产量的对比。如果与期望值没有离差，这些点将位于斜率为 1 的对角直线上。任何一点与对角线的垂直距离都是特殊配合力连同杂交组合产量的抽样误差（如果系群不是高度近交的，在估计特殊配合力时也会有误差，这个误差是由系群中基因型的抽样引起的）。

为例证目的，依据假设系群是来自系群群体的一个随机抽样来计算一般和特殊配合力的方差，尽管事实上它们不是随机选择的。估计组分的方差分析给在表 15-4。*GCA* 和 *SCA* 的平方和是从表 15-3 的数据计算来的；误差均方是论文中报告的数值。杂种真实均值的方差，即扣除误差方差之后，其 54.5%是由一般配合力引起的，45.5%是由特殊配合力引起的。归因于特殊配合力的比例很大，因为这些系群是一个经过选择的群体，因此系群间一般配合力的变异低于在一个随机样本中预测的。

表 15-4　方差分析

来源	自由度	平方和	均方	期望均方
GCA	9	2 179	242.11	$\sigma_E^2 + \sigma_{SCA}^2 + 8\ \sigma_{GCA}^2$
SCA	35	1 617	46.20	$\sigma_E^2 + \sigma_{SCA}^2$
误差			5.36	σ_E^2

$2\ \sigma_{GCA}^2 = 48.98 = 54.5\%$

$\sigma_{SCA}^2 = 40.84 = 45.5\%$

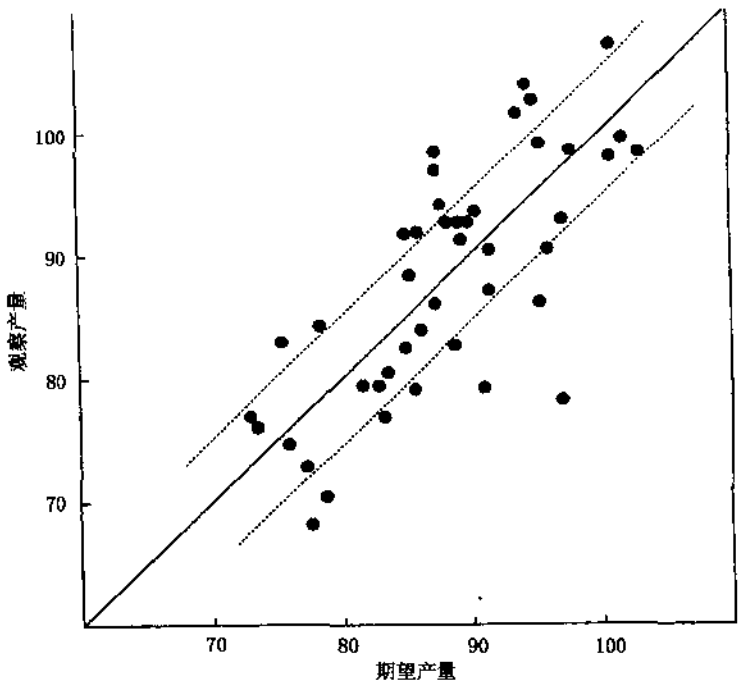


图 15-4　玉米 10 个系群之间杂种的观察和期望产量（每英亩蒲式耳）

每一杂交组合的期望产量是两个亲本系群一般配合力的和，与回归直线的偏差是由特殊配合力和估计杂交组合平均产量的误差引起的。两条点线指出了偏离回归线 ±2 倍标准误的杂种均值位置（数据来自 Sprague 和 Tatum，1942）。

习　题

15.1　果蝇（*Drosophila*）腹部刚毛数的遗传方差主要是加性的（见表 8-4）。假定一个大群体被分成重复系群，每一系群由 20 个亲本繁殖而来，这些系群继续这样下去，直到计算的近交系数达到 50%。如果基础群体的表型方差是 4.0，遗传力是 52%，那么表型方差在系群之间和系群内的组分将是多少？（1）系群内和（2）不计细分成系群时全群的遗传力将是多大？

［解答序号：75］

15.2　假定习题 15.1 中的果蝇基础群体通过全同胞交配近交了 3 个世代。以下而方

式进行的选择其期望反应将是多大？有很多系群，每一个都组成了一单个全同胞家系。从每一系群随机取出雌雄各半的 4 只果蝇；以从系群中抽样的 4 只果蝇的平均值为基础选择出它们中最好的 5%；然后来自每一个中选系群的 4 只果蝇与它们的同胞交配，对后裔进行测定。

[解答序号：85]

15.3 以胸长作为度量的果蝇身体大小，很少有非加性遗传方差（表 8-4）。假定在一个种群中查明胸长的遗传力为 0.34，该种群通过大量个体之间随机交配来保持，但最初它起始于一单对果蝇。在这最初一对果蝇来源的群体中遗传力的估计值是多大？

[解答序号：95]

15.4 小鼠中骨骼存在微小差异是普遍的。在 C57BL 近交系的亚系中研究了其变种发生的 27 个性状。发现两个特定的亚系在 27 个性状中有 5 个性状不同，并且每一种差异可归因于一个突变步骤。在亚系分离之前，这个品系已通过全同胞交配近交了 40 个或更多世代。分离之后，在研究期间通过全同胞交配，一个亚系进一步近交了 21 个世代，另一个亚系近交了 29 个世代。如果差异确实是由突变引起的，那么根据突变影响性状的位点数目以及它们的突变率，将如何解释这些结果？

数据来自 Deol，M.S. *et al*.（1975）*J. Morph.*，**100**，345 ~ 376。

[解答序号：105]

15.5 单株产量参数的下列估计值是从玉米的一个开放传粉的品种中获得的。

平均值 = 308g；$V_P = 5\,798$；$V_A = 864$；$V_D = 188$

如果这个群体近交时没有任何选择，并且系群进行杂交（1）当近交系数为 50% 时，（2）当近交系数实际上为 100% 时，那么在杂种之间和杂种内表型方差组分将是多少？假定 $V_I = 0$，并通过实验设计消除了杂种间的环境差异。

如果每一杂种的平均值是从 20 个单株测定中估计出来的，那么杂种的观察均值方差将是多大？

数据来自 Gardner，C.O.（1977）pp.475 ~ 489 in Pollak，E. *et al*.（eds），*Proc. Int. Conf. Quantitative Genetics*. Iowa State Univ.，Ames，Iowa，USA。

[解答序号：115]

15.6 再次考察习题 15.5 中的杂种。如果在每一阶段都进行 50 次杂交，那么你预测 50 次杂交之间最高的观察平均产量将是多少？然后如果这个最好的杂种重复，来自它的一组新的植株被种植和测定，假定其产量的第一次和第二次计算之间不存在环境差异，你预测平均产量将是多少？

[解答序号：125]

15.7 在正常是自花传粉的金黄色菜豆（*Phaseolus aureus*）的 5 个品种之间进行没有反交的一个双列杂交。杂交组合的每一植株以克为单位的平均产量给在下表。这些品种标明为 A 到 E，它们的产量也给出了（对角线上），虽然对这个问题这些值是不需要的。计算（1）每一品种的一般配合力，（2）与每一杂交组合期望值的离差，即特殊配合力 + 误差。像图 15-4 那样画一散点图。

	A	B	C	D	E
A	9.7	14.1	22.8	16.9	31.8
B	—	3.3	16.5	6.2	12.4
C	—	—	9.0	8.3	9.2
D	—	—	—	6.8	13.1
E	—	—	—	—	12.5

数据来自 Singh，K.B.& Jain，R.P.（1971）*Theor. Appl. Genet.*，**41**，279～281。

[解答序号：135]

第 16 章　近交和杂交：Ⅲ. 应用

近交系杂交产生的杂种在一些植物的改良中起着主要的作用，而最显著的则是玉米。虽然由于近交衰退导致生殖力的严重损失而使得家畜的高度近交系不能有用，但杂交还是广泛应用于动物育种中。因此，动物在轻度近交系之间或不同品种之间进行杂交。本章将阐述为动植物改良应用近交和杂交的原理。我们将主要讨论异交植物，但我们也将用独立的章节来简要考察动物和天然自体受精植物。这里将不提供技术细节；对这些读者应当查阅植物育种的教科书，如 Simmonds（1979）所著。这里我们将作两点简化。首先，假定植物育种中优良性状的唯一标准是产量，虽然在实践中其他性状也必须和产量一起考虑。其次，将不讨论由基因型×环境互作引起的复杂性。显然，改良的杂种必须在与不同年份和不同地区相联系的各种不同环境中表现良好。我们假定这个要求包括在优良性状的评估中。高度近交系的杂交也被用作植物和实验动物遗传分析的一种方法。Mather 和 Jinks（1977，1982）全面描述了对杂交种及其衍生后代的这种分析，在此不赘述。

我们在异交生物体中应用杂交的目的，当然是要产生优良的杂交种或 F_1 个体。首先考虑都起源于同一随机交配基础群体的一组系群。于是杂交种不仅必须优于近交系，而且也要优于这些系群起源的杂交群体。因为杂种优势是超过近交系的优势，因此要寻找比杂种优势更多的某些东西。在第 14 章已指出，当系群未经选择而近交时，则它们所有杂种的平均值预计等于它们起源的杂种群体的平均值。因此单纯进行近交和杂交不能产生任何改良；如果要获得任何改良，必须在某一阶段施加选择。

可是，用来杂交的系群经常起源于不同的基础群体。于是杂种的一些优越性则来自杂种优势。如果两个基础群体的基因频率不同，那么它们之间的杂种将表现出杂种优势，这一点在第 14 章中已作过阐述。同样，起源于两个基础群体未经选择的多组近交系的杂种平均值将高于两个基础群体的平均值。因此即使没有任何选择也能获得一些改良。就动物而言，杂交的系群已经存在，但不存在它们起源的基础群体。因此从杂交中获得的所有增益都是杂种优势，并且唯一的选择是挑选哪一个系群或品种用于杂交。对第 14 章给出的杂种优势的说明，这里不作任何补充。因此我们关注的是令植物实现最大改良的选择。为简化起见，我们假定用于杂交的所有系群都起源于同一基础群体。

从自然选择的效果中可期望获得一些改良。自然选择淘汰了近交过程中致死的和严重有害的基因，就这些基因对有益基因的影响而言，预计杂种平均值会高于基础群体的平均值。但是这种改良不会很大，因为淘汰的有害基因在基础群体中频率低，基因越有害，其频率越低，因此它们对群体平均值的影响很小。以有关位点数及其突变率的假设为基础，人们曾计算过，从隐性有害基因的淘汰中预期获得的适合度改良的最大值只有 5%（Crow，1948，1952）。因此大部分的改良必定来自对有益经济性状实施的人工选择。应用这种人工选择的方法是要考虑的主要题目。可是，也必须同时考虑另一个问题。培育近交系并对它们的杂种进行评估是一个漫长而又昂贵的过程。这种方法超出对原始异交群体实

施简易选择的优点是什么？现在可以部分地回答这个问题 。

近交系的杂交并不能产生基础群体中不存在的基因型。但是，既然最佳基因型只存在于基础群体的特定个体中，那么在特定杂交的每一个体中这些基因型便得以复制。这个方法的主要优点就在于有利基因型的这种复制。当找到一种理想的杂交时，通过在连续世代中重复这种杂交，便可以在需要的任何数目的个体中产生它的基因型。而且，F_1 个体的遗传同一性使它们的表型一致，这种表型一致性就是一种经济利益，尤其对机械收获。例如，成熟时间的一致性意味着所有个体可以在同一时间收获。虽然某一杂交的基因型在基础群体的一个个体中也可能找到，但杂交中该基因型的复制有利于我们几乎无误地度量基因型值；而在基础群体中一个个体的基因型值只能通过它的表型值粗略地度量。此外，只要近交系得以长久保持，在杂交中度量的基因型值还可以无限地复制；而在非近交群体中，通过个体选择，仅育种值可以复制。因此与没有近交的选择相比，近交和杂交很可能是一种较好的改良方法，其条件是当性状的大部分遗传方差为非加性时。

配合力的选择

育种者最终是要在所有可获得的近交系之间找出将提供最好杂种的一对近交系。换句话说，选择最终是要应用杂种。在大量的杂种之间通过选择获得的改进数量取决于杂种之间变异的数量并取决于第 11 章描述过的选择强度。如前一章指出的那样，为了获得高选择强度需要巨大数量的杂种，为了获得杂种之间变异的最大值需要高水平的近交系。可是，通过对系群应用某种初步选择可以节省时间和空间。在两个方面可以做到这一点。首先，最后被选中用于杂交的系群本身必须是能近交繁殖的。因此系群首先被选择的是它们自身的表现。一个近交系的表现与其在杂种中的表现相关，某种程度上这取决于由加性基因引起的方差是多少。可是，这个相关相当小，玉米产量的大约为 0.1（Gama 和 Hallauer，1977），预计依近交表现选择近交系，杂种的改良不能很大。其次，正如前一章阐明的，一个杂种的值由两部分组成：两个亲本系群的一般和特殊配合力。前一章作过概述，系群的一般配合力可以用多种方法来测定，不必做全部可能的杂交。而且，从尚未充分近交的品系也能取得一般配合力的有用指南。

对系群一般配合力的初步选择所获得的改良来自基础群体的加性方差。利用非加性遗传方差所获得的任何进一步的改良一定来自对特殊配合力的选择。通过初步测验没有选择系群的方法；必须通过杂交从杂种中选择最好的系群。因为特殊配合力方差与近交系数的平方或较高的幂成比例（等式 [15.8]），那么从选择特殊配合力中获得很大益处之前，系群必须达到高水平的近交。

一般和特殊配合力的相对重要性 有多少改良预计是来自一般配合力，又有多少是来自特殊配合力呢？如果应用于两者的选择强度相同，那么归因于每一种配合力改良的相对数量将与它们的方差成比例。如果系群是完全近交的，那么一般配合力方差就等于基础群体的加性方差，特殊配合力方差就等于非加性方差（等式 [15.8]）。因此如果基础群体的方差已知，就可以预测来自两种配合力改良的相对数量。在玉米的一个自由授粉的品种中，产量的加性方差占总方差的比例是 0.149，非加性方差则为 0.032（Gardner，1977）。因此加性方差占总遗传方差的比例为 $0.149/0.181 = 0.823$。因此如果近交和杂交应用于这个群体，并且选择应用于杂种，那么可以预计大约 80% 的改良将来自一般配合力，只有

大约 20% 来自特殊配合力。如果这些方差组分是玉米群体特有的，那么似乎是，杂种玉米中绝大部分的改良来自一般配合力并最终来自基础群体的加性方差。这提出了一个问题：在一个没有近交的随机交配群体中通过选择其产量能提高多少。几个实验已表明对自由授粉品种选择是有效的。一个实验中，14 个世代中群体都有反应，产量总改进为 40%（Gardner，1977）。改良的自由授粉品种，它不具有近交种——杂交种重要特征——均匀性，但对基础群体进行事先选择是增加随后从它们中筛选近交系的一般配合力的一种有效方法。

综合群体 当近交系已经产生并对它们的配合力进行了选择，那么这些系群的杂种不会获得进一步的改良。要获得进一步的改良，必须从改良的基础群体中产生一组新的近交系。新的基础群可以通过不经近交的选择得以改良，或者从选择的近交系中得以构建。对许多选择的近交系进行杂交并允许 F_1 代和以后各代随机异花授粉便可以产生一个新的综合群体。在综合群体中，基于加性方差的得以改良的系群的一般配合力得以保持。于是 F_2 代及以后各代的分离有利于产生一组新的近交系，它们的基因组合不同于用来构建群体的系群。通过这种方法，可以获得配合力的进一步改良。目前使用的杂种玉米就是两次或两次以上这样的近交和杂交循环的产品。除了杂种得以改良外，作为近交种的系群的产量也得以提高，这一点具有经济重要性，因为用作商业种植而出售的杂种种子必须由一个近交种亲本产生。

三元和四元杂交；回交

与近交系的低生产率相联系的实际困难可以通过使用三元和四元杂交来克服，虽然在杂交中均匀性和生产性能有一定的损失。在上面提及的改良近交种有用之前，这些杂交一直广泛应用于杂种玉米的生产中。*在三元杂交中*，用两个系群的 F_1 作雌性或种子亲本，需要它有高产性，然后 F_1 与第三个系群杂交。*在四元杂交或双交中*，不同系群对的两个 F_1 代进行杂交。如果从不同基础群体中获得的系群是可用的，那么为了利用群体间杂种优势，则最后的杂交可在不同来源的系群间进行。例如，如果系群 A 和 B 来自一个起源，系群 P 和 Q 来自另一个起源，进行的四元杂交如下：A × B 和 P × Q，紧接着是 AB × PQ。如果不存在上位互作，则从组成系群的单交的表现中能够可靠预测三元和四元杂交的表现。考察三元杂交（A × B）× P 这一实例。让这些字母代表相应系群在一单个位点上携带的等位基因。于是 A × B 的 F_1 具有基因型 AB，当它与系群 P 杂交时以相等比例产生两种基因型 AP 和 BP。这两种基因型分别是系群 A × P 和 B × P 杂交的基因型。因此这两种杂种的平均表现可预测三元杂交的表现。同样，四元杂交（A × B）×（P × Q）的表现可通过 A × P、A × Q、B × P 和 B × Q 的平均值来预测。可是，如果考察不止一个位点，那么在最终杂种中 F_1 亲本分离产生的基因型不会在所用系群配制的任何单交种中出现。因此如果存在上位互作，则单交种不能准确地预测最终杂交种。

由任一特殊的三元或四元杂交所产生的群体都是几种基因型的混合物，所有这些基因型原理上都可由单交产生，但预计在下列情形下三元和四元杂交不同于单交：（1）如果用于杂交的系群经过了选择，并且如果它们的单交种继面产生的任何优越性是上位互作引起的，那么这种优越性中的一些将在三元和四元杂交种中消失。（2）杂种内存在遗传变异并因此丧失了表型一致性。（3）杂种之间的方差减小了，因此最好的三元或四元杂交种并不

像最好的单交种一样好。玉米中例证这些结果的三元和四元杂交与单交的实验比较，见 Weatherspoon（1970）、Otsuka、Eberhart 和 Russell（1972）。

避免近交系低产性的另一种方法是回交。这里只包含两个系群，F_1 与第一次杂交中使用的系群之一再交配，即（A × B）× B。回交后裔中基因型是比例相等的 AB 和 BB。因此，在不存在上位互作的情况下，回交的平均值等于 F_1 和第二次杂交使用的系群的平均值。因此回交的杂种优势不及三元或四元杂交的高。

动物中的杂交　杂交广泛应用于动物生产中，产肉的大多数动物是三元杂交或回交的后裔。前已提及，用于杂交的系群并不是有意近交的。可是，先前已对这些系群的有利性状进行了选择，从而这些系群变成了轻度近交系，一些有利性状因此而减少，尤其是生殖力。杂交的目的，部分是要利用杂种优势来提高生殖力，部分是要结合系群以前被选择过的不同特性。对肉品生产来说，最终产品中一个期望特性是快速生长，最终杂交中期望的是产生大量快速生长的个体。这需要母亲具有良好的生殖力并结合后裔具有良好的生长速度。因此，第一次杂交（A × B）是用来产生具有良好生殖力的 F_1，良好的生殖力来自杂种优势。这些 F_1（AB）个体作为母本与具有良好生长速度的第三个系群（C）杂交来产生（A × B）× C 后裔。或者如果找不到适用的第三个系群，F_1 可与第一次杂交使用的系群之一进行回交。最后的后裔其生长的改良部分来自杂种优势，部分来自公畜系群的加性效应。生长速度不可能总是达到系群的最好水平，但由能生育的 AB 母本所产生的后裔数目增加使杂交有经济优势。Dickerson（1969）和 Sheridan（1981）综述了来自不同杂交类型的增益。

例 16.1

用绵羊品种的一个三元杂交来说明将生殖力杂种优势与公畜系群的优势生长相结合而获得的增益。数据来自 Sidwell、Everson 和 Terrill（1962，1964）。表 16-1 给出了纯种和三元杂交种 3 个性状的表现：（a）生殖力：每只配种母羊的断奶羔羊数；（b）生长速度：断奶时每只羔羊的体重；（c）重要经济性状，每只配种母羊的断奶羔羊总重量，它是（a）×（b）的乘积。杂种的断奶重并不比公畜系群本身的重，但由 F_1 母羊生产出较大数量的断奶羔羊，使三元杂交在断奶羔羊总重量方面优越于最好的纯种。

表 16-1　绵羊纯种和三元杂交的三个性状表现

	每只配种母羊生产的		
	(a) 断奶 羔羊数	(b) 每只羔羊的 断奶重（kg）	(c) 断奶 总重（kg）
纯种			
A = 雪洛泊夏羊（Shropshire）	0.80	23.0	18.4
B = 南丘羊（Southdown）	0.79	19.1	15.1
C = 汉普夏羊（Hampshire）	1.00	29.2	29.2
中亲（$\frac{1}{4}$A + $\frac{1}{4}$B + $\frac{1}{2}$C）	0.90	25.1	22.6
三元杂交（A × B）× C	1.25	27.5	34.4
杂种优势，超过中亲的%	39	10	52
超过最好品种的优势（%）	+25	−6	+18

正反交反复选择（RRS）

杂交种的特殊配合力在该杂交种没配制和测验之前，是无法测定的。因此要对特殊配合力实施合理的高选择强度，必须配制和检测大量的杂交种。难道没有捷径可寻吗？难道通过选择也不能将优越配合力构入系群吗？从第 14 章解释的杂种优势原因来看，很清楚，需要的是影响性状以及表现显性的所有位点上的基因频率存在很大不同的一对系群。因此我们有可能通过选择将两个系群基因频率的这些差异积聚起来。基因频率的差异不是由近交的随机过程产生，而是由选择的定向过程产生的，这既有效又经济。而且，可以同时选择一般和特殊配合力。以这种方法选择配合力称为正反交反复选择（reciprocal recurrent selection）。其理论基础已由 Comstock、Robinson 和 Harvey（1949）以及 Dickerson（1952）检查过。该方法要点如下。

从制备两个群体开始，最好是它们杂交时已知能产生某种杂种优势。这样两个配合力要改良的群体，将被称为系群 A 和 B。进行正反交，大量的 A♂♂与 B♀♀交配，大量的 B♂♂与 A♀♀交配。然后对杂种后裔度量要改良的性状，并且根据其后裔的表现来判断亲本。选择最好的亲本，其余的连同只用来检测亲本配合力的所有杂种后裔全部淘汰。然后中选的个体必须与它们自己系群的成员再交配以产生要被检测的下一世代的亲本。像以前一样，这些个体再杂交，并重复循环进行。应避免有意的近交，因为基因频率的随机变化是不想要的。

一个基本的前提条件是开始的两个系群之间的基因频率应当有一些差异，否则配合力的选择不能引起系群的分化。两个系群中基因频率相同的任一位点将处于平衡，尽管这是一种不稳定的平衡。任何方向的任何改变都将引起有作用的选择，从而增加差异。系群之间的起始差异可以从开始选择已杂交好的两个不同的品种中获得，或者通过故意近交，直到大约为 25%，再依赖于基因频率的随机分化来获得。

关于正反交反复选择实用价值的证据是有争议的。一些商业家禽育种者将其应用于蛋生产中（见 Krosigk 等，1973），玉米中已经产生了有希望的结果（Eberhart，1977）。但另一方面，在家禽和实验动物中正反交反复选择与其他选择方法的直接比较并不是激动人心的（见 Calhoon 和 Bohren，1974；McNew 和 Bell，1976）。

超显性

近交和杂交是否是比没有近交的选择更好的一种改良方法，此问题依有关基因属性的超显性而定。第 2 章中我们曾将适合度的超显性作为保持多态性的一种机制来讨论，并对由连锁引起的真实超显性和拟超显性的不同方式作了阐述。这里我们关注的是有待改良的性状的超显性，但在实践中，超显性是如何产生的并不重要。正如我们已经看出的，两种改良方法都包含选择，因此最重要的区别就在于杂交。固定了不同等位基因的两个系群杂交，产生的 F_1 代的所有个体都是杂合体；并且这是产生全是杂合体的个体组群的唯一方法。在非近交群体中，就一对特定的等位基因而言，杂合体个体数不超过 50%。因此，如果一对特定等位基因的杂合体在优点方面优越于纯合体，那么，与没有近交的选择相比，近交和杂交将是一种更好的改良方法。而且，只有当有关的理想性状或性状的组合存在超显性时，近交和杂交才能获得没有近交的选择所不能获得的优越性。在任何其他显性

条件下，最好的基因型是纯合子之一，理论上所有个体通过选择都能变成纯合的，并没有与近交伴随而来的缺点，并且比依赖于杂交的方法要简单得多。本章前已叙及，当存在较大的非加性遗传方差而加性方差很小时，近交和杂交的潜力最大。现在我们看到这仅是事实的部分：理论上，并不考虑所有实际因素，仅仅是当有关基因至少存在一定程度的超显性时，近交和杂交才超越没有近交的选择。

植物和动物中的许多实验研究，前面各章中已提到一些，都表明所研究性状的超显性并不是基因的一个重要特性。即使在近交和杂交已取得相当成功的玉米产量中这也是真实的（Eberhart，1977）。因此，似乎没有理论理由使人们相信近交和杂交是增加期望性状平均值的一种较好方法。其优势主要在于一致性而不是较高的平均值。

自然自体受精植物

当自体受精植物进行杂交时经常表现出少量的杂种优势。其原因可能是，尽管对少量的有害突变基因施加了不利选择，但它们还是被近交固定下来了。但是，与异交植物表现的优势相比，这种杂种优势要小得多。例如，小麦和大麦（近交种）的变种之间杂交的杂种优势总计大约为 10%（Geiger，1988），而玉米（杂交种）近交系之间杂交的杂种优势通常约为 150%。可是，在商业种植中利用杂种优势并不容易，因为进行杂交在技术上通常很难，并且要产生杂种种子需要在每一世代里大规模地进行杂交。有几种方法可以克服杂交的技术困难，但这些方法仅在几种作物上获得了成功（见 Simmonds，1977，p.231）。因此杂交的目的经常是产生分离。在进行某一杂交之后，允许 F_1 和以后的各代自体受精，以产生一组新的近交系，这些近交系因重组和随机漂变而变得有所分化。目的是要从这些重组系群中找出比两个亲本系群都好的一个或多个系群。来自任一特定杂交的预期改良数量可以从 F_2 的加性遗传方差以及将应用于重组系群的选择强度来预测（见 Jinks 和 Pooni，1976；Pooni 和 Jinks，1979）。

习　题

16.1　如果你将利用习题 15.7 中提到的品种进行三元杂交和四元杂交，为了获得最高的预测产量，你将选择哪些品种，怎样进行杂交？预测的产量是多少？

［解答序号：96］

16.2　对两个品种或品系 A 和 B 进行如下的“轮回杂交”，这里 X 总是指杂种世代。

（1）$A \times B$

（2）$X_1 \times A$

（3）$X_2 \times B$

（4）$X_3 \times A$

假设不存在上位互作和母体效应，根据纯种和单交种的表现来计算 1～4 代中各杂种后代的预期生产表现。

［解答序号：106］

16.3　三个品种 A、B、C 间的轮回杂交如下进行，这里 X 仍指杂种后代。

（1）$A \times B$

（2）$X_1 \times C$

(3) $X_2 \times A$

(4) $X_3 \times B$

(5) $X_4 \times C$

假设没有上位和母体效应，根据纯种和单交种的表现来计算每一世代的预期生产表现。

[解答序号：116]

16.4 假设在轮回杂交中使用的品系的所有单交种表现出相同数量的杂种优势。那么在轮回杂交的连续世代中这种单交种的杂种优势将达到多大比例？对习题 16.2 中两系轮回杂交 4 个世代和习题 16.3 中三系轮回杂交的 5 个世代求解这个问题。

[解答序号：126]

16.5 当将轮回杂交育种应用于动物中时，用于杂交的母本总是杂种而父本总是纯种。于是系统常具有有用的特征：纯种本身需要生产的雌性并不比更新所需要的多。以下数据是 3 个品种切斯特白猪（C）、杜洛克猪（D）和约克夏猪（Y）以及它们的单交种中个体猪 154 日龄的平均体重（W）和 154 日龄每窝猪的头数（N）。轮回杂交由（Y♀ × C♂）♀ × D♂开始，根据习题 16.3 的解答计算轮回杂交的 5 个世代中每一世代每窝猪的预期平均总重。作如下的简化假设：每窝猪的头数只取决于母本基因型，体重则只取决于个体基因型，上位是不重要的。

	C	D	Y	CD	CY	DY
W (kg)	78	88	84	92	90	96
N	6.6	6.3	7.9	7.4	8.2	7.3

数据来自 Schneider, J.F. (1978) Ph. D. Thesis, Iowa State Univ., Ames, Iowa, USA。

[解答序号：136]

第 17 章　尺　　度

前几章已多次提到了对度量数量性状的适宜尺度的选择。可是，对尺度选择包括什么的解释以及对适宜标准的讨论延缓到这一部分，因为直到理解了从数据中获得的推论性质，我们才能正确地评价这些问题。换句话说，尺度的选择必须与数据的用途联系起来。来自任何实验或实际研究的数据都是以对性状度量最方便的形式获得的。也就是说，表型值是以克、厘米、天、数字或最方便的任何度量单位来记录的。争论的要点是这些原始数据在它们接受分析或解释之前是否应当转换成另一种尺度。尺度转换意味着把原始单位转换成对数、倒数或其他一些函数，并按数据使用的目的来判断什么样的转换最适宜。

假若每个性状都有自己的“固有”尺度，那么在性状中生物学过程按尺度表达，尺度有效，这是诱人的。因此生长是一种几何的而非算术的过程，几何尺度将显得最自然。例如，体重为 20g 的小鼠增重 1g 与体重为 2g 的小鼠增重 1g 没有相同的生物学意义；但是，增重 10% 对两者来说有大致相同的意义。因此，对体重的度量转换成对数似乎是合适的。可是，这主要是主观判断，需要一些客观标准用于尺度的选择。然而，不同的标准在它们表示的尺度中经常不一致，而且，应用于同一性状的同一标准在不同的群体中可能表示不同的尺度。因此，每个性状必须有其“固有”的和正确的尺度的想法多半是虚幻的。

概括地说，做尺度转换有 3 个主要原因：①使得分布成为正态；②使得方差独立于平均值；③消除或减少非加性互作。每一种情形下，尺度的选择标准都是达到特定目标的一种经验。当需要尺度转换但未做时，一定现象发生了，被称为*尺度效应*（scale effects），当做了适当的转换时，尺度效应就消失了。上面指出的目标完全等于对消除这些尺度效应的说明。我们将特别讨论将算术尺度转换为几何尺度的对数转换。这可能是最普遍和最有用的转换。Wright（1968，第 10 章）阐述了其他的转换。参考对数转换列出的一般原理同样适用于其他转换。

分布与方差

首先考察表型值的分布。图 17-1 显示了 3 种分布，似乎是以算术尺度用原始数据绘制的。如果数据先转换成对数，所有 3 种分布将是对称的和正态的。有两个重要的特点需要注意。第一，偏离正态性的程度取决于与平均数有关的变异大小。这可以从不是非常明显不对称的上面两条曲线（a）和（b）与类似的下面一条曲线（c）的比较中看出。变异大小和平均数之间的关系决定偏离正态性的程度，它的最佳表达为变异系数，即标准差与平均数的比率，变异系数通常乘以 100，从而将其转化为百分率的形式。上面两条曲线的变异系数是 20%，而下面一条曲线则为 50%。因此，转换成对数对分布的形状没有产生可察觉的差异，除非变异系数相当高，即大于 20% 左右。因此，当变异系数不高于 20% 时，不以严格的正态分布为基础的统计方法，比如方差分析，能应用于未转换的数据。当变异系数低于 20% 时，转换成其他尺度也就没有什么必要了。

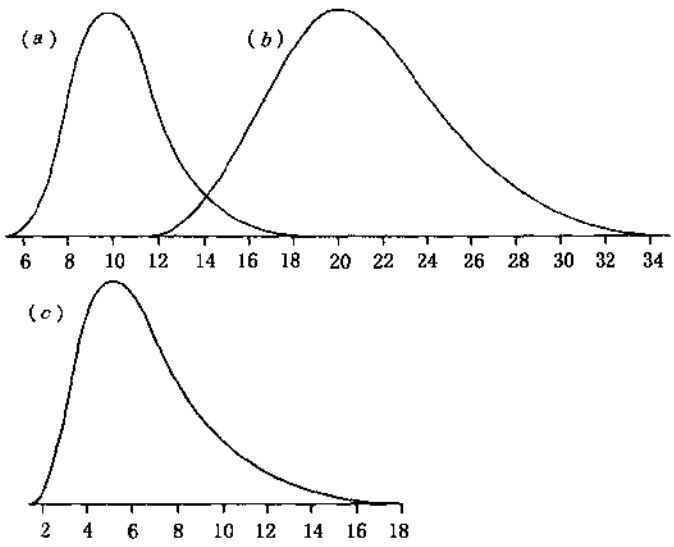


图 17-1 以算术尺度作图表示的对数尺度的对称和正态的分布

在图 17-1 中需要注意的第二点是方差，以算术单位计算时，方差随平均数增加而增加。这在曲线（a）和（b）中可以看出。在对数单位中，这两条曲线有相同的方差，但平均数不同。（b）的平均数，或严格地说是众数，是（a）的两倍，而且算术单位中标准差相应地也是两倍。虽然分布不是非常显著偏斜，并且也没有显示出强烈需要转换，但是由于平均数的不同，方差也有很大的不同。于是，这里是一种最普遍的尺度效应，即方差随群体平均数的变化而变化。如果性状是身体大小一类用重量单位度量的，那么图 17-1 中的两条曲线（a）和（b）可以很好地代表经若干世代双向选择而歧化的两个群体。通常能找到这类性状，当它们的平均数增加时，其方差也增加，当它们的平均数减少时，其方差也减少。图 17-2 显示了小鼠实验的一个例子，性状是 60 日龄体重。应注意的是单独考虑的这 3 种分布，没有一种是充分的不对称和非正态分布，需要根据这个标准做尺度转换；但为了使方差独立于平均数，很显然需要转换。

如果诸如上面讨论的方差变化的现象在度量适当转换时便消失了，换句话说，如果它们变化的原因能归于度量尺度，那么这种现象就称为尺度效应。虽然将这种现象给予尺度效应的标笺，或者经转换而消除之，但它们仍然是真实的。例如，当以克来度量体重时，体重大的鼠确实比体重小的鼠更容易发生变异（表 17-1）。承认这种现象是一种尺度效应的获益是，不必为作出解释而深究性状的遗传特性。

我们结合图 17-1 中的曲线（a）和（b）指出的标准差与平均数比率为对数转换的适合性提供一个方便的检验。如果两种分布按对数尺度有相同的方差，那么以算术单位表示的变异系数将是相同的。因此，变异系数的稳定性表示了对数尺度的方差稳定性。而且，如果要比较方差，我们就可以简单地比较变异系数来代替以对数单位来表示的方差。图 17-2 显示的分布及其标准差和变异系数给出数字说明。虽然变异系数不一样，但比标准差更为相似，这表明由选择引起的方差变化至少大部分可归因于度量尺度。

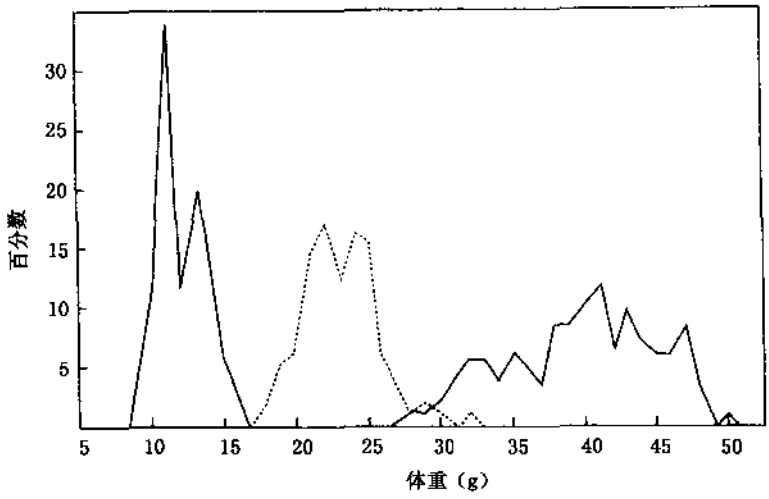


图 17-2 60 日龄雄性小鼠体重的分布

中心：选择前的基础群体。左边和右边：双向选择 21 世代之后的小和大的品系（根据 MacArthur，1949）。

表 17-1 不同群体体重的平均数、标准差和变异系数

	体重小的群体	未选择群体	体重大的群体
平均数	11.97	23.16	39.85
标准差	1.71	2.56	5.10
变异系数（%）	14.3	11.1	12.8

当需要对数转换时，并不总是每个个体的度量都需要转换。采用下面的公式可以方便地进行平均数和方差的转换（Wright，1968，p.229）：

$$(\overline{\log x}) = \log\bar{x} - \frac{1}{2}\log(1 + C^2) \qquad [17.1]$$

$$\sigma^2_{(\log x)} = 0.4343\log(1 + C^2) \qquad [17.2]$$

第一步将算术均数转换为对数均数，第二步将从算术值计算的方差转换为对数方差。这些公式中 C 是从算术值中计算得来的、以 $\sigma/\bar{x}$ 的形式表示的变异系数，对数以 10 为底。只有依对数尺度分布确实是正态的，这些公式才是精确的。

当关于方差的结论严格地取决于消除任何尺度效应时，找出方差或标准差与平均数之间的经验关系可能是必要的。如果具有不同平均数的几个群体是可用的，并且如果它们的方差不同不存在尺度效应之外的其他原因，则可以做到这一点。于是关联标准差与平均数的回归方程给出了具有一特定平均数的另一个群体的期望标准差。如果回归是线性的，则回归方程是 $\sigma = a + b\bar{x}$，其中 σ 是平均数为 $\bar{x}$ 的群体的期望标准差，a 是截距，b 是回归系数。方差独立于平均数的尺度是 $X = \log(x + a/b)$，其中 x 是原始度量值，X 是它的转换值（Wright，1968，p.232）。

让我们回到选择的结果并进一步追究之。如果方差随着作为选择结果的平均数的变化而变化，那么选择差和选择反应也是这样。因此正像我们已考察过的性状那样，每世代的选择反应将随着向上选择的进展而预期增加，在向下选择中则相应地减少。于是双向选择的反应将是不对称的。最可能以这种方式归因于尺度效应的一个不对称反应实例显示在图17-3中。以算术单位绘制成图，在（a）中，向上选择的反应比向下选择的要大得多。转换成对数，显示在（b）中，致使反应更加接近对称。这样做并没有消除这个事实，即按照度量，在选择下性状增加的比它下降的要多得多。但它解释了不对称性，不需要精细的假设。从选择反应的图形表示中看出，消除尺度效应的一个方便的方法是以现实遗传力来绘制选择反应图，正如在第11章中解释的和图11-5中所说明的那样。现实遗传力是选择反应与选择差的比率，极少受尺度效应的影响（Falconer，1954）。

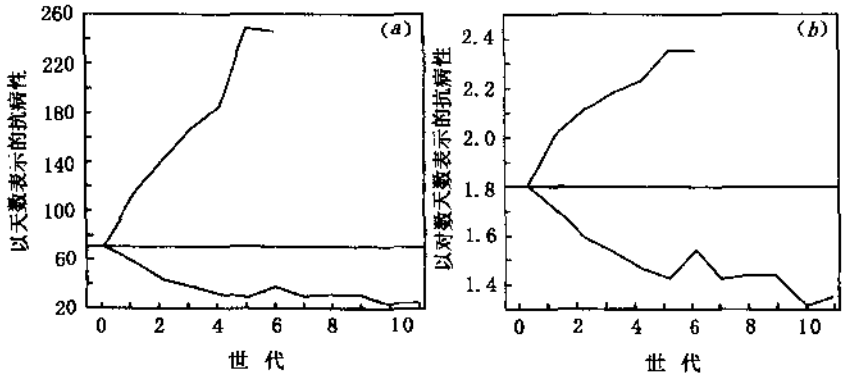


图17-3 对大鼠龋齿抗性双向选择的反应

抗性用天数度量，以算术尺度绘制在（a）中，以对数尺度绘制在（b）中。通过公式［17.1］将算术平均数转换为对数平均数。变异系数很高，大约50%，大致保持不变。向上选择没有向下选择进行的世代多，简单的原因是抗性提高延长了世代间隔（数据来自Hunt，Hoppert和Erwin，1944）。

互 作

现在我们转向什么可能是尺度转换的更为基本的效应——它对遗传方差的表观性质的影响。为了理解这个问题，我们必须回到一单个位点并考察基因的效应或作用方式。试想一个位点有两个等位基因，其作用方式是几何的，A_2A_2 的基因型值比 A_1A_2 的大50%，并且 A_1A_2 的基因型值也比 A_1A_1 的大50%。因此，按照对数尺度就没有显性存在，因为杂合子恰好处子两种纯合子的中间。现在假设基因型值是以算术单位比如克度量的，并且 A_1A_1 的基因型值是10个单位。于是 A_1A_2 将是15个单位，A_2A_2 将是22.5个单位。因此，按照算术尺度，A_1 对 A_2 为部分显性，因而杂合子不再介于两种纯合子的中间。这样显性度受到了度量尺度的影响，显性方差的数量比例也如此。可是，尺度转换的这种效应通常相当小。引起基因型值之间50%差异的基因，正像已考察过的那样，将是一个主效基因，很容易逐一地识别出来。但即使这样，算术尺度的显性度并不很大。可能具有1%或10%效应的微效基因其显性几乎不受影响。

如同显性受尺度影响的方式一样，不同位点间的上位互作也受尺度的影响。如果以对数单位度量基因型值，那么具有几何效应的位点没有互作地联合起来。但当以算术单位度

量时，将有归因于上位互作的离差。因此，互作方差的大小也受度量尺度的影响。下面的例子说明了互作对尺度的依赖性。

例 17.1

小鼠的侏儒基因是影响身体大小的一个主效基因，纯合体身体大小减小很多。这个基因的效应是在不同的遗传背景下研究的（King，1955）。通过重复回交，将出现在小身体中选品系中的这个基因转移至大身体中选品系中。在两个品系中以及转移期间，度量了侏儒型纯合子和正常个体（即杂合体加上正常纯合体）之间的平均差异，对相同窝中的侏儒型和正常个体间进行比较。结果如图 17-4 所示。侏儒型和正常个体间的差异随着正常个体的体重增加而增加。在小品系的背景下，侏儒型个体大约比正常个体小 7g，但是在大品系的背景下，侏儒型个体大约要小 14g。因此侏儒基因和影响身体大小的其他基因之间存在上位互作。但是如果侏儒基因的效应用比率表示，则它是常量并独立于现存的其他基因。无论其实际体重是多少，侏儒个体大约是它们正常窝同伴体重的一半。因此，如果以对数单位进行比较，就没有上位互作。

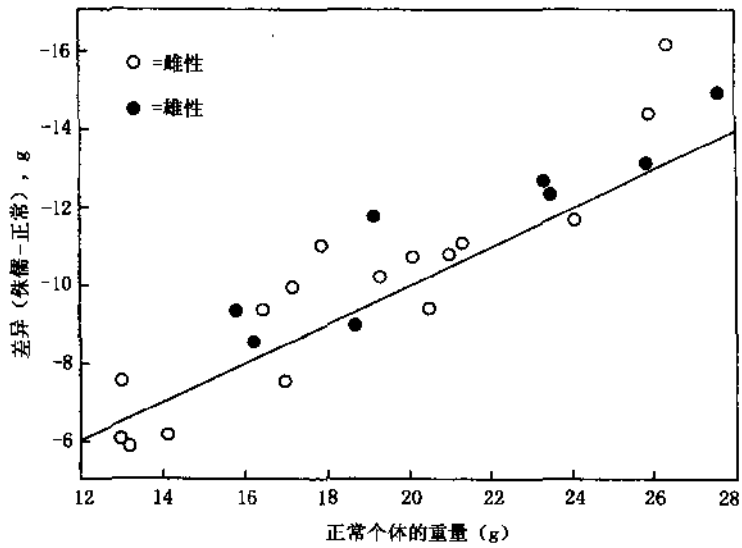


图 17-4　不同体重小鼠中侏儒基因的效应

侏儒型和它们正常的窝同伴之间 6 周龄体重的差异以同一窝中正常个体的体重为基础绘制成图。直线不是配合的回归线，而是表示这种关系：侏儒个体体重 = 0.5 × 正常个体体重。见例 17.1（数据来自 King，1955）。

因此通常情况下，尺度转换可以消除或减少归因于上位互作的方差，于是这个方差可以标记为尺度效应。如果要从其有效性取决于没有互作的一个分析中得出结论，那么消除或减少互作方差的尺度转换可能是有用的。基因型和环境之间的互作也可由尺度效应产生，转换对于消除或减少这些互作可能是有用的。然而，上位或基因型 × 环境的互作，不可能总是通过尺度转换来消除或甚至减少。例如，没有意义的转换能消除引起次序颠倒的互作，就像图 8-2 所说明的。

有两种方法可以找到消除互作的适宜尺度。第一种方法是通过比较特定因素的效应，从而找到使效应成为加性的一种尺度。上面例 17.1 例证了这一点。在不同遗传背景下比较一单个基因的效应，可以发现根据对数尺度，在所有遗传背景下这个基因都增加同样的数量。令效应变为加性的这个特定因素是环境的而不是遗传的。当具有不同平均数的两个群体是有用的并且它们能够杂交时，可以采用适宜尺度的第二种检验方法消除上位互作。正如第 14 章和第 16 章所指出的，在不存在上位互作的情况下，F_2 和回交世代的平均数期望如下：

$$\left.\begin{aligned} F_2 &= \frac{1}{2}(F_1 + \overline{P}) \\ B_1 &= \frac{1}{2}(F_1 + P_1) \\ B_2 &= \frac{1}{2}(F_1 + P_2) \end{aligned}\right\} \qquad [17.3]$$

其中，$\overline{P}$ 是双亲群体平均数，所有其他符号都是相应世代的平均数。选择的尺度要使观察平均数与它们的期望值最接近。详细细节见 Mather 和 Jinks（1982，p.71）。

结 论

这一章我们概述了最常见的一些尺度效应，并且指出了尺度转换可能有助于结果解释和结论获得的环境。可是，没有充分的理由就不应当做尺度转换。实验观察的第一个目的是描述群体的遗传特性，尺度转换模糊了而不是阐明了这种描述。例如，如果发现存在上位，这是描述的一个不可缺少的部分，并且标记为上位比标记为尺度效应更好。尺度转换基本上是一个统计手段，其目的在于简化数据分析或使得从分析中得出有效结论成为可能。尺度转换有时也有助于结果的解释。例如，如果转换成对数尺度时发现上位效应消失了，那么我们可以得出结论：不同位点的效应是通过相乘而不是相加结合起来的。或者，如果方差的一个差异有充分的理由归因于尺度效应，那么我们就不必寻求较复杂的遗传解释。可是，尺度选择产生了对有关结果解释的麻烦问题。尺度转换的合乎逻辑的理由只能来自某一标准而不是要获得的结论特性。如果没有独立的标准，则争论会变得周而复始，并且尺度效应和其他一些解释之间的区别就变得毫无意义。还有一个更为基本的困难：适合一个群体的尺度可能不适合另一个群体，并且适合于变异的遗传组分和环境组分的尺度可能不同。对番茄变种间大量杂交种的“每室重量”性状的分析有力地例证了这个困难（Powers，1950）。根据同一标准——分布的正态性，发现这个性状在一些杂种中需要算术尺度，而在另一些杂种中则需要几何尺度；而且，在一些杂种的 F_2 世代中，遗传变异需要一种尺度而环境方差需要另一种尺度。

习 题

17.1 图 17-2 指出的数据，如果期望 3 个系群的方差相等，则需要转换成对数。那么对数体重的标准差将是多少？在选择反应中数据表现出相当大的不对称性。通过对数转换能消除这种不对称吗？

[解答序号：10]

17.2 下面的数据是小鼠不同体重的选择品系以及这些品系杂种的六周龄体重（g），

有3个“大小组”，即大的、对照和小的。在每一个大小组中有6个重复系群（图12-1涉及这些系群）。在同一大小组的系群之间以及不同大小组的系群之间进行杂交。目的是要查明不同大小的系群间的杂种其杂种优势是否比同一大小组的系群内杂种的要大。表中第一行给出了每一大小组中重复的平均体重，表中其余各行给出了正反交平均起来的平均体重。通过将平均数转换成对数，关于杂种优势的结论将作怎样的修改？

	大	对照	小
系群均值	28.74	20.99	14.91
大	30.68	26.00	21.85
对照	—	21.91	18.48
小	—	—	14.84

数据来自 Kumar, C.K.B.（1980）爱丁堡大学博士学位论文。

［解答序号：20］

17.3 下表指出了具有X染色体和常染色体不同组合的雄果蝇（*Drosophila*）的腹侧刚毛平均数。有两种X染色体，一种来自高刚毛数系群，另一种来自低刚毛数系群。这两种X染色体被放到来自不同水平刚毛数的系群的常染色体背景中。X染色体和常染色体之间存在强烈的上位互作。你能找到一种尺度转换，消除互作并使X染色体的效应在所有背景中都相同吗？这不是容易做的。当平均数很低时除外，知道有4根刚毛，每边两根，比其余的要大，结构上也不同并且从不缺失是有帮助的。

常染色体背景	X染色体来源		差异（高－低）
	高	低	
A	9.49	7.75	1.74
B	13.34	10.84	2.50
C	22.72	16.36	6.36
D	34.87	24.84	10.03
E	47.80	32.80	15.00

数据来自 McPhee, C.P. & Robertson, A.（1970）*Genet. Res.*, **16**, 1～16。

［解答序号：30］

第 18 章　阈性状

有很多性状具有生物学意义或经济价值，它们以一种不连续的方式变化，但并不以简单的孟德尔方式遗传。熟悉的例子是对疾病的敏感性，有两个表型类别——感染或未感染；还有较大家畜的窝产仔数，它们通常一次产生一个幼畜，但有时是 2 个或 3 个。也有不连续的解剖学差异，例如小鼠的椎骨数，它的遗传学已被广泛研究。这类性状乍看起来好像在数量遗传学领域之外，但对它们进行遗传分析时，发现它们与连续变化的性状按相同的方式遗传。

趋向和阈值

了解这类性状遗传性的线索在于这样的认识，即性状具有潜在连续性，这种连续性有一个阈值（threshold），因此，这便强加给可见表现一种不连续性，如图 18-1 所示。当潜在变量低于这个阈值水平时，个体有一种形式的表型表现，比如说是“正常”表型，当它高于该阈值时，个体表现为另一种表型，例如，“感染的”。如果讨论的阈性状为人类疾病时，这个潜在连续变量便被称为趋向（liability），这里仍将使用这个术语。趋向的连续变异起源上既是遗传的又是环境的，可以是某种底物的浓度，或某一事物发育过程的速度，也就是说，在原理上可用常规方法将它们作为数量性状来度量和研究。它可能是几种不同生理或发育过程的复合物，但没有必要了解这些过程是如何结合起来产生这种趋向的，甚至不必了解它们真正是什么。

通过对窝产仔数的考察，可正确判断潜在变量的认识是实际的。小鼠或猪的窝产仔数，尽管实际上是明显不连续的，但可作为连续变量来处理，因为有足够大的数目等级。母牛的窝产仔数只有两类，单胎和双胎，两个犊牛以上极其稀少。但没有理由认为牛生双胞胎的生理原因不同于小鼠或猪窝产仔数的。这两种情形中潜在变量主要由循环促性腺激素的水平组成，该激素决定排出的卵数以及影响胚胎存活的子宫内因子，而且对牛来说，还影响决定单卵双生的因子。

两种类别，一个阈值

让我们首先考察只有两种表型类别并用一单个阈值区分它们的性状。这两个类别将被定为正常的和感染的。在表型水平或可见尺度上，个体只有两个可能的值，指定 0 表示正常，1 表示感染。可是，按照被感染个体的比例或百分率，个体的群体，例如家系或作为一个整体的群体，会有任何数值。这被称为发生率（incidence），或在人类疾病范围中称为流行率（prevalence）。作为群体或组群的一种简单描述，发生率是很恰当的，但表示发生率的百分率尺度对很多用途是不恰当的，因为在百分率尺度上方差随平均值而不同。因此对遗传分析，发生率必须被转换成平均趋向。为了作出这个转换，必须定义趋向是正态分布的。这个定义要求，如果我们直接度量趋向，那么通过某种尺度转换使它成为正态分

布是可能的。以后我们将回到这个要求；同时我们假定这个条件可以满足。由于趋向是正态分布的，于是趋向的单位是它的标准差 σ。因此平均趋向通过单尾正态离差 x 与发生率相关联，x 是阈值偏离平均值的离差，以趋向的标准差单位表示。不同发生率的 x 值列于附录表 A 中。

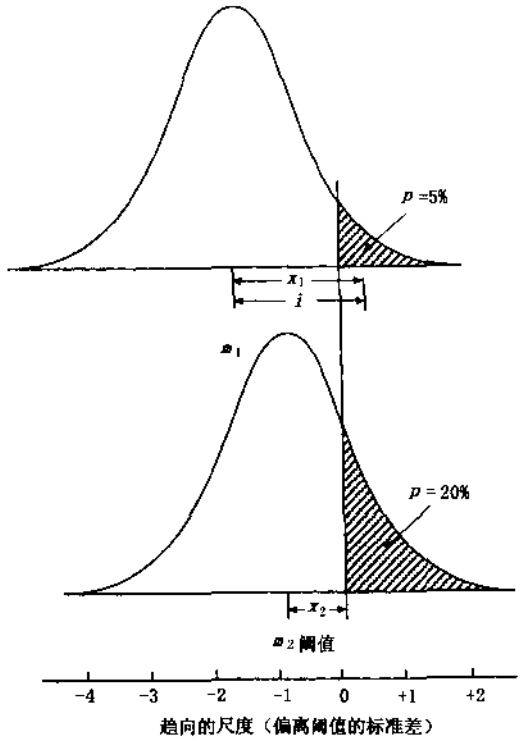


图 18-1　阈性状发生率不同，因而平均趋向不同的两个群体或组群

两组群的趋向方差相同，平均值相差达 0.8 个趋向标准差。x 是阈值偏离平均值的正态离差；i 是受感染个体偏离它们组群均值的平均值离差。

平均值的比较　考察具有不同发生率的两个群体或组群，如图 18-1 所示。它们的平均趋向相差多少？为了比较不同的组群，阈值必须被定义为固定的；也就是说，所有组群中阈值处于相同水平的趋向。于是任何一组群的平均值都用阈值离差表示，单位为 σ。换句话说，阈值被认为是趋向尺度上的起源或零点。图 18-1 中上面的组群有一个发生率 $p_1 = 0.05$，它给出 $x_1 = 1.6\ \sigma_1$。这样平均趋向是 $m_1 = -1.6\ \sigma_1$（必须注意平均趋向的符号。如果阈值高于平均值，那么平均值便低于阈值因而是负的）。类似地，图 18-1 中下面的组群具有一个发生率 0.20 且平均值 $m_2 = -0.8\ \sigma_2$。注意到平均值是以它们自己群体的标准差为单位表示的。我们不能进一步进行平均值的比较，除非我们假定两个标准差相同。如果这能被接受为一个有理由的假设，那么 $\sigma_1 = \sigma_2 = \sigma$，且 $m_2 - m_1 = 0.8\ \sigma$；两组

群的平均值相差趋向 0.8 个标准差。

趋向的遗传力 假定图 18-1 上图的分布代表亲本世代，从中选择了感染个体作为亲本。当这些亲本繁殖时，它们产生具有下图分布的后裔。知道亲本世代和后裔的发生率，我们就有了计算后裔对趋向的中亲值回归所需要的一切，并且可以从回归中计算趋向的遗传力。首先考察选择反应，这是亲本和后裔世代之间平均值的差异，上面已给出为 0.8 σ，并假设两个世代的方差是相等的。实际上方差不会完全相同，但暂时忽略引入的小误差。现在考察选择差。这是亲本世代中被感染个体的平均趋向，作为一种与它们群体平均值的离差。用作亲本的个体比例可能低于发生率；换句话说，不是所有的感染个体都用作亲本。但只要所有亲本是感染的，则可预料使用的那些个体的平均值就与所有感染个体的平均值相同。以标准差单位表示的感染个体的平均值就等于选择强度 i，相当于作为中选比例的发生率。i 的值列于附录表 A 中。对发生率 $p=0.05$ 来说，选择强度 $i=2.1$，因此选择差是 $S=2.1\ \sigma$（见等式［11.5］）。后裔对中亲值的回归是反应与选择差的比率（等式［11.1］），则 $R/S=0.8\ \sigma/(2.1\ \sigma)=0.38$。最后，如果后裔和其亲本之间没有环境相似性，则趋向的遗传力就等于后裔对中亲值的回归。

上面说明的有关亲本和后裔的回归和遗传力的计算可适用于任何种类的亲缘关系。假定图 18-1 的上图分布代表一个群体，下图分布代表任何特定的感染个体亲属种类，例如全同胞。于是计算的回归是一个个体对它的亲属的。因为两个组群——群体和亲属的方差是大致相等的，所以回归就近似地等于相关。因此，对图 18-1 的发生率来说，就趋向而言全同胞的相关就等于 0.38。在由共同环境引起的相似性以及显性都缺乏的情况下，则遗传力将被估计为全同胞相关的 2 倍，即 0.76。

上面说明的计算可汇总在下列公式中。任何特定种类亲属之间的趋向相关由下式给出：

$$t=\frac{m_R-m_P}{i}=\frac{x_P-x_R}{i} \qquad [18.1]$$

其中，下标 P 和 R 分别指群体和亲属，m 是阈值离差的平均值，x 是阈值偏离平均值的正态离差，i 是感染个体偏离群体平均值的平均离差。符号适合于感染个体比正常个体具有较高趋向的尺度。于是与等式［10.5］一样，从相关中获得的遗传力为：

$$h^2=t/r \qquad [18.2]$$

其中，r 是表 9-3 给出的亲缘系数。对第一级亲属、第二级亲属和第三级亲属，r 分别为 $\frac{1}{2}$、$\frac{1}{4}$和$\frac{1}{8}$。当亲属是受感染双亲的后裔时，合适的“r”是对中亲的回归，它是 1。

平均趋向即正态离差 x 的抽样方差是：

$$\sigma_x^2=\frac{1-p}{i^2A}$$

其中，A 是感染的数目，p 是发生率。由等式［18.1］估计相关抽样方差是相当复杂的（见 Falconer，1965a），但如果能假定群体发生率误差可忽略，则抽样方差可由下式给出：

$$\sigma_t^2=\frac{1-p_R}{i_P^2 i_R^2 A_R} \qquad \text{（近似）}$$

其中，下标 P 和 R 分别指群体和亲属，A 是感染的数目。

通过假定感染个体的亲属方差与作为一个整体的群体方差相同而引入的误差导致由等式［18.1］估计的相关过低，要低一个因子 5% 或 10%。计算不等方差的一个修改公式如下（Reich，James 和 Morris，1972）：

$$t=\frac{x-x_R\sqrt{[1-(x^2-x_R^2)(1-(x/i))]}}{i+x_R^2(i-x)} \qquad [18.3]$$

其中，没有下标 x 和 i 指的是群体，x_R 指的是亲属；取平方根的符号令 t 在 0 和 1 之间。

例 18.1

隐睾病是在一些猪群中公猪发生的一种先天性缺陷。表中数据指的是 Mikami 和 Fredeen（1979）报道的一个猪群。猪群即群体的发生率是 3.9%，患病公猪的全同胞之间的发生率是 11.6%。从附录表 A 中获得的 x 和 i 对应值给在表 18-1。通过等式［18.1］计算的全同胞相关如下：

$$t=\frac{1.762-1.195}{2.165}=0.26$$

假定没有共同环境也没有显性，由等式［18.2］得出的趋向遗传力为：

$$h^2=2t=0.52$$

用等式［18.3］中更准确的公式计算，相关为 $t=0.28$，给出 $h^2=0.56$。

表 18-1　公猪隐睾病的发生率及其对应的 x 和 i 值

	数　目		发生率		
	患病	总数	p%	x	i
群体（P）	44	1 129	3.9	1.762	2.165
全同胞（R）	25	215	11.6	1.195	

趋向遗传力的计算被广泛应用于人类疾病的遗传研究中。通过患特定疾病的可疑病人收集有关他们亲属患病状况的资料。于是可按照上面的方法计算有关趋向的相关。但是，对有关遗传力相关的解释易受共同环境引起的相似性不确定性的影响，如在第 10 章强调的那样。假定没有环境相似性，获得的趋向遗传力估计值从精神分裂症的 85% 变化到先天心脏病的 35%（Emery，1976，p.54）。就计算家族中复发危险率而言，遗传力的知识在遗传咨询中是有用的，因为它有利于正确地结合有关家族的所有信息。关于应用于人类疾病的进一步的细节，见 Falconer（1965a，1967），Curnow 和 Smith（1975）。

趋向模型的充分性

作为一种正态分布的连续变量，趋向的定义必须经更严密地检测。正如先前注意到的，它暗示了这一假设：结合起来给出趋向的潜在变量可通过尺度转换成为正态。反过来这暗示着趋向的分布是单峰的。实际上如果它是双峰或多峰的，便没有合适的尺度转换可使它的分布成为正态。于是相关和遗传力的计算将无效。有两种主要方式可引起双峰或三峰分布：第一，如果存在一单个基因，它对趋向的影响与相关的残差变异比较是相当大的，第二，如果存在一种环境因子，它的影响相对于其他变异更大。影响疾病趋向的这种

环境因子可以暴露给病原体。因此，只有趋向是多因子的，这意味着变异原因很多，它们的效应都相对小，并且由在多于一个或几个位点上的基因进行遗传控制，只有这样，根据趋向进行的遗传分析才是有效的。

没有方法事先了解有效分析的这些要求是否满足，因为不能度量趋向来看看它的分布是否是单峰的。可是，人们可以看看获得的结果是否是合理的和一致的。例如，超过100%的遗传力显然是不能接受的，表明存在一单个主效基因。而且，在没有由共同环境引起的相似性的情况下，从不同种类亲属中估计的遗传力应当是相同的。总的说来，获得的结果具有合理的一致性并且不会产生怀疑趋向模型充分性的强烈理由。用来测验不同种类亲属的一致性，同时检测单个主效基因以及环境相似性（如果存在）的分析方法被称为复合分离分析（见 Morton 和 MacLean，1974；Lalouel 等，1977），但是这里描述它是太复杂了。可对具有两个阈值的性状进行充分性差异检测的一种方法，将在下面描述。

尺度关系

给一个阈性状的两种表型类别指定武断值 0 和 1，同时计算与这些数值有关的亲属间的相关，这都是可能的。要做到这一点偏向于采用表型表现作为度量趋向的一种很粗糙的刻度仪器；事实上，该仪器仅有一个刻度记号。如果要估计方差组分，这将引人大量的度量误差，它显然是环境方差。在（0，1）尺度上由度量误差引起的方差的大小取决于发生率；发生率为 0.5 时它最小，而发生率较低或较高时它变得较大。因此，根据（0，1）尺度计算的相关随发生率而变化，随发生率低于或高于 0.5，它都变小。像上面描述的那样，转换成趋向尺度消除了由度量误差引起的方差，并导致从它们中获得的相关和遗传力独立于发生率。通过一种简单的关系，相关或遗传力能从一种尺度转换成其他尺度（Dempster 和 Lerner，1950）。如果 t_c 是在趋向连续尺度上的相关，t_{01}是根据（0，1）尺度计算的同一相关，那么两者通过下式关联：

$$t_c = t_{01}\left(\frac{1-p}{i^2 p}\right) \qquad [18.4]$$

其中，p 是群体中感染个体的发生率，i 是对应的平均趋向。两种遗传力以同一方式关联。已经看出由上面方式获得的遗传力估计值没有由等式［18.3］获得的估计值那么准确（Mercer 和 Hill，1984）。

三种类别，两个阈值

如果有 3 个表型类别，并且就趋向而言类别可逻辑地排序，也就是说，倘若有生物学理由可以相信一个类别的趋向处于其他两个的中间，则阈性状的遗传分析可深入进行。一个实例是单胎、双胞胎和三胞胎。于是区分为三种类别的两个阈值处于趋向的不同水平。这两个阈值标记为趋向尺度上的固定点，它们在所有组群中都是相同的。因此这两个阈值之间趋向的差异提供了与标准差不同的一个固定的趋向单位。这使得比较不同组群的标准差或方差以及比较预计有不同方差的组群平均值成为可能。这种认识最容易用数字例子说明。

考察图 18-2 中例证的两个群体。它们有不同的平均值和不同的方差。阈值 T_1 和 T_2 是趋向尺度上的固定点，并且它们之间的差异是趋向的 1 个阈值单位（t.u.）。趋向的尺

度显示在底部，并且 T_1 的位置上为 0。第一步是要以阈值单位来表示每一个标准差。用发生率可以做到这一点，结果如下。用 p' 表示大于 T_1 的个体比例，即中间和极端类别共同的发生率，用 p'' 表示大于 T_2 的个体比例，即极端类别的发生率。用 x' 和 x'' 表示 T_1 和 T_2 偏离群体均值的离差，以标准差表示之。于是对任何一个群体，阈值之间的差异是 $(x''-x')\sigma$，其中 σ 是群体的标准差。阈值之间的差异定义为 1t.u.，因此 $(x''-x')\sigma=$ 1t.u.，用阈值单位表示的标准差由下式给出：

$$\sigma=\frac{1}{x''-x'}\text{t.u.} \qquad [18.5]$$

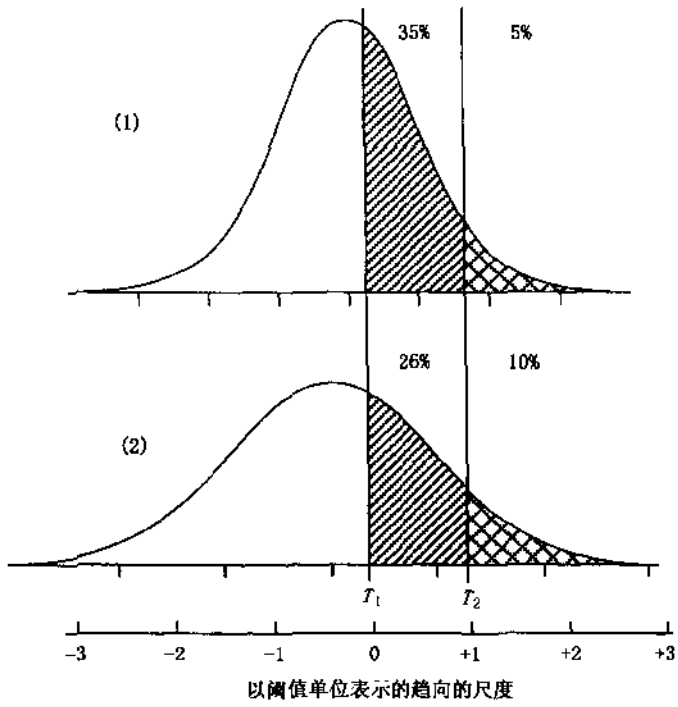


图 18-2　具有 3 个表型类别和两个阈值的阈性状，具有不同平均值和不同方差的两个群体中趋向的分布

图 18-2 中群体标准差的计算如表 18-2 所示，其中 x 值从附录表 A 中获得：

表 18-2　两个群体的标准差（用阈值单位表示）

群体	p'(%)	p''(%)	x'	x''	1t.u.	σ
(1)	40	5	0.25	1.64	1.39 σ_1	0.72t.u.
(2)	36	10	0.36	1.28	0.92 σ_2	1.09t.u.

这样，发现群体（2）的标准差是群体（1）的 1.5 倍。用阈值单位表示的作为 T_1 离差的平均值如下：

$$m_1 = -x_1' \sigma_1 = -0.25 \times 0.72\text{t.u.} = -0.18\text{t.u.}$$

$$m_2 = -x_2' \sigma_2 = -0.36 \times 1.09\text{t.u.} = -0.39\text{t.u.}$$

这样，群体（2）的平均趋向低于群体（1）0.21 个阈值单位。

当用上面描述的方式可以估计方差时，研究具有不同发生率的系群间杂种以及比较 F_1 和 F_2 世代的方差就成为可能。这提供了趋向模型充分性的又一种测验，特别是对处于其他两种类别之间的一种类别趋向的解释。下面的例子说明了小鼠近交系杂种的这样一个研究。亲本系群、F_1、F_2 和两个回交种的平均值和方差与从数量性状中预期的结果非常一致；这种情形下没有理由怀疑阈值模型的有效性。

例 18.2

小鼠的前荐椎数在 25 和 27 之间变化，前荐骨被定义为位于与荐骨交融在一起的第一椎骨的前部。因此这个性状反映了荐骨与脊柱交融在一起的纵向位置。通常只有两个数字出现于任一近交系中，这里我们考察具有 25 和 26 个前荐椎数，但比例极不相同的两个近交系。第三个表型类别由较少个体提供，这些个体是不对称的，一侧有 25 个而另一侧有 26 个前荐骨，仅在一侧通过最后一根椎骨与荐骨交融在一起。显然，对不对称小鼠作为趋向的中间存在某些怀疑；实际上，可能它们在发育中没得到较好的调整。但是通过杂种的遗传分析，将它们作为中间值来看待似乎是有理由的。这里的数据指的是由 Green (1962) 描述的几个杂种之一。品系 C3H 有 13% 的个体具有 26 个椎骨，而品系 C57BL 有 90% 的个体是这样的。表 18-3 给出了不同世代的发生率以及按照上面所计算出的平均值和标准差。平均值是用阈值单位表示的、将“25”和“不对称”区分开来的阈值离差。中亲值以及从等式 [17.3] 计算得来的 F_2 和回交世代的期望平均值也显示在表 18-3 中。总的来说，结果与期望是一致的：F_1 是 P_1 和 P_2 之间的中间值；F_2 接近 F_1 的平均值但方差有较大增加；回交平均值位于 F_1 和亲代平均值之间，并且有 F_1 和 F_2 方差之间的方差。两个亲本系群的方差很不同是唯一的异常。

表 18-3 不同世代小鼠前荐骨数的发生率、平均值和标准差

	小鼠数目	p'%	p''%	m	$\overline{P}$ $E(m)$	σ	$\sigma^2 = V_P$
P_1	282	12.8	8.5	−4.8	−0.7	4.2	18.0
P_2	619	96.4	89.8	+3.4	−0.7	1.9	3.6
F_1	532	50.0	31.8	0.0	—	2.1	4.5
F_2	206	56.8	46.1	+0.6	−0.3	3.7	13.8
B_1	205	33.7	21.5	−1.1	−2.4	2.7	7.4
B_2	194	75.3	60.8	+1.7	+1.7	2.4	5.9

F_2 中遗传决定程度 V_G/V_P 可从 F_2 和 F_1 之间的方差差异中估计出来，结果如下：

$$F_2: \quad V_G + V_E = 13.8$$

$$F_1: \quad V_E = 4.5$$

$$F_2 - F_1: \quad V_G = 9.3$$

$$V_G/(V_G + V_E) = 0.68$$

F_2 中趋向变异的 68% 是可遗传的。而且，这是一个很合理的结果。

从这一章已经说明的来看，很明显阈性状不能提供数量遗传学研究的理想材料，因为用它们所做的遗传分析在范围上有限并且受制于人们不愿作出的假设，除了不得已而为之。如果能找到一个与趋向密切相关的连续变化性状，那么显然将它作为一个数量性状而不是阈性状来分析更好。例如，“存活时间”可用来代替“抗性对感染性”；或者一种畸形可根据严重的程度进行分等。

阈性状的选择

对阈性状实施选择不含有遗传分析的理论困难。它在减少畸形的发生率以及改变实验动物对处理的反应，例如，增加或降低抗药性等，具有实际重要性。我们将考察有两个明显类别的阈性状，并且首先涉及个体选择。

通常情况下，选择反应取决于选择差。对于一个连续变化的性状，选择差基本上与选择比例无关，但由于下列原因它取决于发生率。我们可以单独繁殖那些理想表型类别中的个体，但是我们不能区别具高和低趋向的那些个体。因此中选个体是理想类别中的一个随机样本，它们的平均值就是理想类别中的平均值，而不管我们选择的是理想类别中的全部还是它的一部分。因此选择的比例比发生率还小不会产生任何好处。另一方面，如果选择的比例不得不大于发生率时，则我们将被迫使用非理想类别的一些个体。它们的平均趋向将低于群体平均值，因此对非理想个体将使用某些负向选择（非理想类别的平均值很容易计算出，为 $-ip/(1-p)$，其中 i 是发生率为 p 的理想类别的平均值）。这些考虑使得这一点很清楚，即当理想类别的发生率等于为更新而必须选择的比例时可获得最大的选择差。两者之间差异越大，选择就越没有效果。由于这个原因，对稀有的畸形进行不利个体选择是很少有成效的。作为一个例子，考察具有 5% 发生率的一种畸形。于是 95% 的个体是正常的，它们的一个随机样本有低于群体平均值 0.1 σ 的平均趋向，提供很小的选择差。如果我们对畸形在另一个方向上选择，例如必须从群体中选择 20% 作为亲本，于是从 100 个中选择出的这 20 个个体中，只有 5 个是畸形的，剩下的 15 个必然是正常的。畸形个体每一个贡献差别为 +2.1 σ，正常个体每一个贡献差别为 −0.1 σ，使得净选择差为 0.4 σ。相比之下，对连续变化的性状选择 20% 可产生 1.4 σ 的选择差。

对阈性状进行家系选择比个体选择有效得多，特别是当发生率很低时。不可能很准确地从个体的感染或正常状态中了解趋向尺度上个体的表型。但是，可以非常准确地从感染成员的比例中了解到一个家系的平均趋向。当然，准确度取决于家系数目，也取决于家系的发生率，因为家系代表来自二项分布的一个样本。50% 的发生率给出家系间的最好区别。为了评价不同选择方法的效率，对引用的例 18.1 进行了研究。发现个体选择需要 50 个世代才能将隐睾病的发生率从 5% 降低到 1%，而半同胞家系选择做到这一点预计只需要 3 代。后裔测验是家系选择的一种形式；各种情况下对它的效率的评价见 Curnow (1984)。

正如我们已经看到的，当理想类别的发生率等于中选比例时，个体选择下的选择差最大。在某些情况下，通过外部方法控制发生率以及使发生率更接近最佳是可能的。例如，

如果被选性状是对某种处理的反应，那么处理可加强或减弱以便发生率发生变化。这种改变的发生率可最好视为阈值相对于群体平均趋向的一种变换。如果这种处理随着选择的进程可进一步改变，以保持发生率尽可能地近似等于中选比例，则可获得个体选择的最大反应。通过使群体或它的一部分接受起始的处理可评价取得的进展。

遗传同化　应用环境手段改变阈值这一原理的一种非常有趣的结果，是称为“遗传同化”的现象（Waddington，1953）。如果一个阈性状作为一种环境刺激的结果而出现，并且对这个性状做选择，那么实际上可能使得它自发出现，而无需环境刺激。用这种方式，通过完全正统的选择原理使最初的“获得性状”变成了“遗传性状”（Waddington，1942）。在这类情形中，有两个阈值，一个是自发的，另一个是诱导的，如图 18-3 所示。开始自发的阈值在群体变异范围之外，这样没有表型变异，因而不能应用选择［图 18-3（a）］。可是，诱导的阈值在群体覆盖的趋向范围之内，它允许个体趋向可被选择挑出的分布的一端。用这种方式可改变群体的平均基因型值。如果这种改变进行足够大，最终一些个体将越过自发阈值而作为自发变种出现［图 18-3（b）］。当自发发生率变得足够高时，选择可以继续下去而无需环境刺激的辅助，并且自发的发生率可能会进一步提高［图 18-3（c）］。

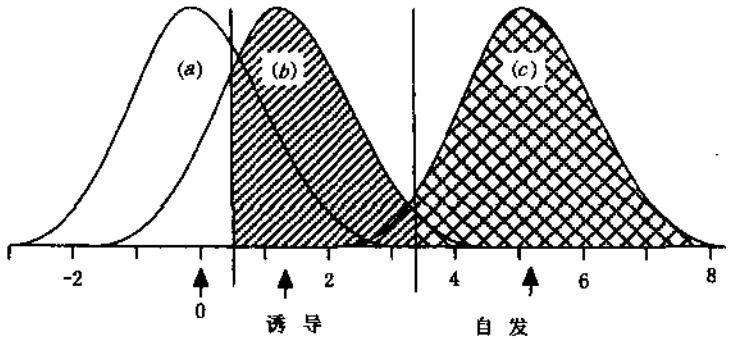


图 18-3　用来说明例 18.3 描述的遗传同化的图解

趋向的分布是以偏离起始群体平均值的标准差标记的。竖线表示诱导和自发阈值的位置，箭头指示下列 3 个选择阶段中的群体平均值。

（a）选择前：　　发生率－诱导＝30%，自发＝0%

（b）某种选择之后：　　发生率－诱导＝80%，自发＝2%

（c）进一步选择之后：　　发生率－诱导＝100%，自发＝95%

例 18.3

Waddington（1953）描述了黑腹果蝇（*Drosophila melanogaster*）中遗传同化的一个实验论证。性状是翅膀后部横脉的缺乏。基础群体中不存在这种畸形的果蝇，但通过对蛹壳进行热休克处理造成大约 30% 的无横脉个体出现。对经处理的果蝇实施两个方向的选择，14 个世代后诱导性状的发生率在上选系群中上升到 80%，而在下选系群中则降至 8%。这时，无横脉果蝇开始在上选系群的未处理果蝇中少量出现，到第 16 世代自发发生率在 1% 和 2% 之间。然后没有处理，选择继续进行，群体被细分成许多系群。经没有进一步处理的选择，4 个最好的系群的自发发生率达 67% ~ 95%。图 18-3 中的分布例证了上选的进展。曲线（b）显示了 2% 的自发发生率和 80% 的诱导发生率，因此大致与第 16 世代

一致。在固定方差的假设下，这个阶段平均值的改变可达 1.36 个标准差。曲线（c）显示了 95% 的自发发生率并且代表最后表现最大进展的那个系群。它的平均趋向比起始群体的要高 5.15 个标准差。

习　题

18.1　（人）髋部先天性脱位的一项家族研究给出了下列的结果。总共研究了 397 名“标记病人”的亲属。第一级亲属大多是父母和全同胞，第二级大多是祖父母、叔父、姨母、侄子和侄女，第三级全是嫡表兄妹。检查的亲属数和受畸形影响的数目如下。

	受影响	总数
第一级	35	1 777
第二级	16	4 746
第三级	8	4 220

整个群体的发生率是每 1 000 人中大约有 1 人。从这些数据中，（近似）评估三类亲属趋向的相关，并估计趋向遗传力。你认为哪一种遗传力估计值可能是最可靠的？

数据来自 Wynne-Davies（1970）pp.316 ~ 338 in Emery, A.E.H.（ed.）, *Modern Trends in Human Genetics*-1. Butterworth, London。

[解答序号：94]

18.2　下面的数据来自瑞典弗里生牛生双胞胎的一项分析，某种程度上简化了。从记录中随机挑选出有一次双生的母牛（或青年母牛）。然后查明了它们的母亲和它们的女儿在第四次、第五次或第六次产犊中具有双胞胎的比例。下面给出了这些发生率以及整个品种中的发生率。在初产具有双胞胎的那些母牛中，随后产犊中具有双胞胎的比例给在下表的最后一行。近似计算产生双胞胎这种趋向的遗传力和重复力。

群体	3.5%
母亲	4.6%
女儿	4.8%
重复产犊	10.0%

数据来自 Johansson, I. *et al*.（1974）*Hereditas*, **78**, 201 ~ 234。

[解答序号：104]

18.3　分析生双胞胎的一种选择方法是将它作为“窝产仔数”来处理，如果个体生单个犊牛则其值为 1，如果它们生双胞胎则其值为 2。当采用这种方法分析时，发现瑞典弗里生牛的半同胞相关为 0.0058。这与习题 18.2 中计算出的趋向遗传力一致吗？

[解答序号：114]

18.4　毛腐烂是澳大利亚美利奴绵羊与潮湿天气相联系的一种损伤体况。可是，在易感性方面存在遗传变异。在发生率为 23% 的一个种群中趋向遗传力被估计为 0.3。降低发生率显然是理想的；增加发生率对实验种群可能是有用的。如果以个体为基础选出 10% 的雄性和 50% 的雌性，进行降低和增加发生率的两个方向选择，那么两个世代后发生率将是多少？

Kevin D. Atkins 热情提供数据。

[解答序号：124]

18.5 在选择小鼠大体格的一个品系中出现了一种多趾（指）（过多的指和趾）。开始它的频率低并且只有后脚受到影响。对受影响个体经过几个世代的繁殖增加了频率并导致了后脚和前脚都受影响的个体出现。这样有三种表型类别：正常（N）、仅后脚受影响（H）、后脚和前脚都受影响（F）。选择 5 个世代之后的频率为：

N	H	F
20%	50%	30%

不同交配类型的后裔为 H 的趋向位于 N 和 F 中间提供了充分的证据。

计算（1）以标准差单位表示的、两个阈值之间的趋向差别，（2）作为区分 N 和 H 的阈值离差，以阈值单位表示的群体平均值，（3）作为群体平均值的离差，以阈值单位表示的三种表型的每一种平均趋向。

数据来自 Roberts，R.C. & Mendell，N.R.（1975）*Genet. Res.*，**191**，427～444。

[解答序号：134]

第 19 章　相关性状

本章论及两个数量性状之间的关系，尤其是在群体中个体的性状值是相关的那些性状，不论是正相关还是负相关。由于下面 3 个主要理由人们对相关性状感兴趣。首先，关于由基因多效性作用引起相关的遗传原因：基因多效性是主效基因的普遍特性，但到目前为止，我们很少有机会考察它在数量遗传学中的影响。其次，关于由选择引起的变化：了解一个性状的改良将如何引起其他性状同时发生变化是重要的。再次，关于自然选择：一个数量性状与适合度之间的关系是确定自然群体中该性状遗传特性的主要力量。可是，这最后一点将在下一章中讨论。

遗传和环境相关

遗传研究中，有必要区别性状之间相关的两个原因，即遗传的和环境的。相关的遗传原因主要是基因多效性，虽然连锁是瞬时相关的一个原因，特别在从歧化品系之间杂交获得的群体中。简单地说基因多效性就是影响两个或更多性状的一个基因特性，因此如果这个基因分离，它就会引起它所影响的性状同时发生变异。例如，提高生长速度的基因既增加体高又增加体重，因此这些基因倾向于引起这两个性状之间的相关。可是，增加膘情的基因影响体重却不影响体高，因此不是相关的原因。由基因多效性引起的相关程度表示了两个性状受相同基因影响的程度。但基因多效性引起的相关是影响两个性状的所有分离基因的全面的或纯的效应。一些基因可能增加两个性状，而其他一些基因增加一个性状并减少另一个性状；前者倾向于引起正相关，后者则引起负相关。因此基因多效性不一定引起一种可检测的相关。在两个性状同样受环境条件的差异影响时，环境是相关的一种原因。同样，由环境原因引起的相关也是所有变化的环境因素的总体效应；一些因素倾向于引起正相关，其他则引起负相关。

两个性状之间可直接观察到的关系是表型值相关，或表型相关。它由群体中大量个体的两个性状的度量值所确定。可是，假定我们不仅知道度量个体性状的表型值，而且知道两个性状的基因型值和环境离差，于是我们能够计算两个性状的基因型值之间的相关和环境离差之间的相关，因此可独立评估相关的遗传和环境原因。另外，如果我们知道个体的育种值，我们还能确定育种值的相关。原则上，显性离差之间和各种互作离差之间也存在相关。论及所有这些相关将复杂得难以处理，但幸运的是，没有必要这样做，因为从两个相关的角度来处理实际问题就很足够了。这就是遗传相关，它是育种值的相关，还有环境相关，严格地讲，环境相关不光是环境离差的相关，而是环境离差和非加性遗传离差一起的相关。换句话说，就像将一个性状的方差分剖成两个组分，即加性遗传组分和所有剩余组分，对很多目的是足够的一样，现在两个性状的协方差也只需分剖成同样的两个组分，这样，“遗传”和“环境”相关与协方差分剖成的加性遗传组分和所有剩余的组分相吻合。估计这两个相关的方法将在后面说明。要考察的第一个问题是遗传和环境相关如何结合在

一起给出可直接观察到的表型相关。

下列符号将在整个这一章中使用：

X 和 Y：所考察的两个性状。

r_P 两个性状 X 和 Y 之间的表型相关。

r_A X 和 Y 之间的遗传相关（即育种值的相关）。

r_E X 和 Y 之间的环境相关（包括非加性遗传效应）。

cov 两个性状 X 和 Y 之间的协方差，下标 P、A、E 与相关具有相同的含义。

σ^2 和 σ 方差和标准差，具有像上面一样的下标 P、A、E 和按照所指的性状具有下标 X 或 Y；例如，σ^2_{PX} = 性状 X 的表型方差。

h^2 遗传力，按照性状具有下标 X 或 Y。

e^2 $= 1 - h^2$。

（符号 r_G 经常用来表示遗传相关，但是，既然所指的相关几乎总是育种值的相关，为了与前面各章保持一致，这里将采用符号 r_A。）

一个相关，无论它的特性是什么，都是合适的协方差与两个标准差乘积的比率。例如，表型相关是：

$$r_P = \frac{\mathrm{cov}_P}{\sigma_{PX}\ \sigma_{PY}}$$

表型协方差可写为：

$$\mathrm{cov}_P = r_P\ \sigma_{PX}\sigma_{PY}$$

表型协方差是遗传与环境协方差之和，即：

$$\mathrm{cov}_P = \mathrm{cov}_A + \mathrm{cov}_E$$

像上面一样，用相关和标准差的形式书写这些协方差，则给出：

$$r_P\ \sigma_{PX}\sigma_{PY} = r_A\ \sigma_{AX}\sigma_{AY} + r_E\ \sigma_{EX}\sigma_{EY}$$

现在注意到 $\sigma_A = h\ \sigma_P$ 和 $\sigma_E = e\ \sigma_P$。代入这些则给出：

$$r_P\ \sigma_{PX}\sigma_{PY} = r_A h_X\ \sigma_{PX} h_Y\ \sigma_{PY} + r_E\ e_X\ \sigma_{PX} e_Y\ \sigma_{PY}$$

两边同除以 $\sigma_{PX}\ \sigma_{PY}$，导致：

$$r_P = h_X h_Y r_A + e_X e_Y r_E \qquad [19.1]$$

这看出相关的遗传和环境原因如何结合在一起给出表型相关。如果两个性状遗传力低，那么表型相关主要由环境相关确定；如果它们的遗传力高，那么遗传相关更重要。表型相关的两重性使得这一点很明显：即不能仅仅从表型相关来确定遗传相关的大小乃至其符号。

表 19-1 给出了遗传和环境相关的一些例子。在某些情形中，遗传和环境相关在大小上不同，或甚至符号相反。在另外一些情形中，两种相关具有相同的符号和没有很大差异的大小，这是更经常的情况。大的差异，尤其是符号的差异表明变异的遗传和环境来源通过不同的生理学机制影响性状。

遗传相关表示两种度量反映遗传上属同一性状的程度。例如，果蝇（*Drosophila*）两翅的长度显然一定是同一性状（翅长）的度量。但翅长和胸长，虽然两者都是身体大小的

度量，但确实不是同一性状；它们之间的遗传相关大约是 0.75（Reeve 和 Robertson，1953）。就此而言，遗传相关对多次度量的重复力的解释有影响。在第 8 章已经说过，要是不同度量针对的是同一个遗传性状，则重复力具有准确的遗传解释，对此无需解释。现在可以明白这种要求的意义是度量之间的遗传相关一定为 1。如果等式［19.1］中两个性状 X 和 Y 是同一性状，那么表型相关的表达式可简化成 $r_P = h^2 + e^2$，这与等式［8.12］中的重复力是等价的，因为这里 e^2 是一般环境引起的方差（V_{Eg}）和非加性遗传方差的比例。即使遗传相关经常在某种程度上低于 1，重复力仍然是一个有用的概念。

表 19-1　表型、遗传和环境相关的一些例子

	r_P	r_A	r_E
人（Grundbacher，1974）			
血清免疫球蛋白含量，IgG：IgM	0.20	0.07	0.31
牛（Barker 和 Robertson，1966）			
产奶量：乳脂率（第 1 泌乳期）	−0.26	−0.38	−0.18
第 1：第 2 泌乳期产奶量	0.40	0.75	0.26
猪（Smith，King 和 Gilbert，1962）			
增重：背膘厚	0.00	0.13	−0.18
增重：效率	0.66	0.69	0.64
禽（Emsley，Dickerson 和 Kashyap，1977）			
体重：蛋重	0.33	0.42	0.23
体重：产蛋量	0.01	−0.17	0.08
蛋重：产蛋量	−0.05	−0.31	0.02
小鼠（Rutledge，Eisen 和 Legates，1973）			
体重：尾长	0.45	0.29	0.56
黑腹果蝇（*Drosophila melanogaster*）（Sheridan 等，1968）			
腹部：腹侧刚毛数	0.14	0.41	0.06

注：引用的估计值系指特定环境中的特定群体；不应当认为它们是普遍适用的。

有一些配对性状，它们之间不存在表型相关。这些是在同一个体上不能进行两次度量的性状。公畜和母畜性成熟的年龄就是两个这类性状的例子。虽然不能测定出表型相关，但这两个性状在遗传和环境上可能仍然存在相关，这两种相关能够估计出来。

遗传相关的估计

遗传相关的估计以亲属之间的相似性为基础，以类似于第 10 章描述的估计遗传力的方法进行估计。因此这里只需描述这个方法的原理而不是详细的细节。我们不是根据方差分析来计算一个性状的方差组分，而是根据与方差分析具有完全相同形式的协方差分析来计算两个性状的协方差组分。我们不是从个体值的平方以及根据变异来源对平方和进行分剖入手，而是从每一个体中两个性状值的乘积以及根据变异来源对乘积和的分剖入手。这导致协方差观察组分的估计，关于协方差有因组分的解释与表 10-7 给出的方差组分的解释完全相同。因此，在半同胞家系分析中，公畜之间的协方差组分估计为 $\frac{1}{4}\mathrm{cov}_A$，即是两个性状育种值协方差的 $\frac{1}{4}$。对相关估计每个性状的方差组分也是必需的。这样公畜间方差

组分估计为$\frac{1}{4}\sigma^2_{AX}$和$\frac{1}{4}\sigma^2_{AY}$。因此可获得遗传相关为：

$$r_A = \frac{\mathrm{cov}_{XY}}{\sqrt{(\mathrm{var}_X \mathrm{var}_Y)}} \qquad [19.2]$$

其中，var 和 cov 表示方差和协方差组分。

也能用亲子关系来估计遗传相关。为从后代和亲本之间相似性中估计一个性状的遗传力，我们取亲本或中亲值与后代平均值的乘积计算该性状的亲子协方差。为估计两个性状之间的遗传相关，我们计算称谓“交叉方差”的值，它从亲本 X 性状值和后代 Y 性状值的乘积中获得。这个“交叉方差”是两个性状遗传协方差的一半，即$\frac{1}{2}\mathrm{cov}_A$。每一个性状各自的亲子协方差也是必需的，于是遗传相关可由下式给出：

$$r_A = \frac{\mathrm{cov}_{XY}}{\sqrt{(\mathrm{cov}_{XX}\mathrm{cov}_{YY})}} \qquad [19.3]$$

其中，cov_{XY}是“交叉方差”，cov_{XX}和 cov_{YY}是每一个性状各自的亲子协方差。交叉方差可由亲本 X 值和后代 Y 值或由亲本 Y 值和后代 X 值来计算。如果两个都有效，那么应当采用算术平均值。采用一种类似于现实遗传力估计的方法，也可以从选择反应中估计遗传相关。这将在下一部分解释。

提供遗传相关估计值的数据也提供了相关性状的遗传力和表型相关的估计值。然后根据等式［19.1］就能得到环境相关。如果使用的是高度近交系，那么环境相关可直接由系内，或者更好由系间杂交的 F_1 的表型相关来估计。

遗传相关的估计值经常有相当大的抽样误差，因而很少能非常准确。而且，遗传相关受基因频率的强烈影响（Bohren，Hill 和 Robertson，1966），因此它们在不同群体中可能显著不同。由于这些原因，必须认为表 19-1 中引用的例子是近似值，对其他群体不一定正确。遗传相关的抽样方差是一个复杂的问题。（详细细节见 Van Vleck 和 Henderson，1961；Hammond 和 Nicholas，1972。）从 Reeve（1955*b*）和 Robertson（1959*b*）推导的标准误的一个近似公式是：

$$\sigma_{(r_A)} = \frac{1 - r_A^2}{\sqrt{2}} \sqrt{\frac{\sigma_{(h_X^2)}\ \sigma_{(h_Y^2)}}{h_X^2 h_Y^2}} \qquad [19.4]$$

其中，σ 表示标准误。这个公式指的是以两个交叉协方差为基础的估计值；如果只采用一个交叉协方差，则$\sqrt{2}$将被忽略。因为两个遗传力的标准误出现在分子上，所以采用第 10 章描述的方法令遗传力估计值的抽样方差最小化的实验设计对遗传相关的估计也是最优设计。

选择的相关反应

考察的下一个问题涉及选择反应：如果我们选择性状 X，那么相关性状 Y 的变化将是怎样呢？当选择应用于另一个性状 X，一个性状 Y 的期望反应可用下面方法推导。性状 X 即直接选择的性状的反应等于被选个体的平均育种值。第 11 章对此有解释。因此性状 Y 随之发生的变化可通过 Y 的育种值对 X 的育种值的回归给出。这个回归是：

$$b_{(A)YX} = \frac{\text{cov}_A}{\sigma^2_{AX}} = r_A \frac{\sigma_{AY}}{\sigma_{AX}}$$

根据等式［11.4］，直接选择的性状 X 的反应是：

$$R_X = ih_X \ \sigma_{AX}$$

因此，性状 Y 的相关反应是：

$$CR_Y = b_{(A)YX} R_X \qquad [19.5a]$$

$$= ih_X \ \sigma_{AX} r_A \frac{\sigma_{AY}}{\sigma_{AX}}$$

$$= ih_X r_A \ \sigma_{AY} \qquad [19.5b]$$

或者代入 $\sigma_{AY} = h_Y \sigma_{PY}$，则相关反应变为：

$$CR_Y = ih_X h_Y r_A \ \sigma_{PY} \qquad [19.6]$$

$h_X h_Y r_A$ 这一项被称为协遗传力，因为它等价于直接选择反应 $R = ih^2_X \sigma_{PX}$（等式［11.3］）中的 h^2。

因此，如果两个性状的遗传相关和遗传力已知，就能够预测相关性状的反应。相反，如果通过实验测定了相关反应，并且两个遗传力已知，那么可估计出遗传相关。如果性状 Y 的遗传力是从选择反应中估计出来的现实遗传力，那么有必要进行双重选择实验。在一个系中对性状 X 进行选择，在另一个系中对性状 Y 进行选择。然后每一个性状的直接和相关两种反应都能测得。这种类型的实验提供了遗传相关的两个估计值（根据等式［19.6］），每一个性状的相关反应提供一个估值；如果等式［19.6］表示的相关反应理论充分地描述了观察反应，那么两个估计值应当一致。采用下列公式，无需遗传力估值，便可从这样的双重选择试验中取得遗传相关的联合估计值：

$$r^2_A = \frac{CR_X}{R_X} \frac{CR_Y}{R_Y} \qquad [19.7]$$

这一公式很容易从等式［11.4］和等式［19.5b］中推导。

例 19.1

在黑腹果蝇（*Drosophila melanogaster*）的翅长和胸长的研究中，Reeve 和 Robertson（1953）从选择反应中估计了身体大小的这两种度量之间的遗传相关。有两对选择系；一对做增加和减少胸长的选择，另一对做增加和减少翅长的选择。在每一个系中，都测定了不直接选择的性状的相关反应以及直接选择的性状的反应。根据等式［19.7］获得了遗传相关的两个估计值，一个来自上选的反应，另一个来自下选的反应。另外，在未选择群体中从亲子协方差和全同胞协方差中也获得了遗传相关的估计值。这 4 个估计值如表 19-2 所示。

表 19-2　果蝇翅长和胸长的遗传相关

方法	遗传相关
亲子	0.74
全同胞	0.75
上选	0.71
下选	0.73

来自选择的估计值与来自未选择群体的估计值之间的一致性表明相关反应与从未选择群体的遗传分析中所预测的非常接近。

我们不能总是期望，像上面例子所示那样，观察和预测的相关反应之间具有紧密的一致性，实际上，这种情况并不常见，尤其是在遗传相关低的情况下。而且，双重选择实验给出的遗传相关估计值经常是不一致的。有两个原因可引起低的预测能力和相关反应的不一致。第一个原因是基础群体中遗传相关估计值的低准确性，这是由已提到的大抽样误差引起。第二个原因是遗传相关对基因频率变化的敏感性（Bohren，Hill 和 Robertson，1966）；因为选择本身和随机漂变的影响，遗传相关进而相关反应在选择过程中会迅速变化。由于这些原因，在实践中应用相关反应理论必然缺乏一些可靠性。可是，我们将在下一部分略微深入探索这一理论的实践启示，但必须注意的是在实践中依靠这一理论不可能总是取得好的效果。

相关选择差 对一个性状 X 进行选择时，任一表型相关性状 Y 均因此而具有相关的或表观的选择差。换句话说，选择过 X 性状的个体，其 Y 性状的平均值将不同于群体平均值。乍看起来，似乎可以利用这种相关选择差来预测相关反应或者采用类似于等式［11.7］（$h^2 = R/S$）的方法估计相关性状的遗传力。可是，不幸的是，相关选择差对这两个目的没有任何用处。理由简要如下。设 S_Y'表示 Y 的相关选择差。于是 $S_Y' = b_{(P)YX}S_X$。采用等式［19.5a］的形式书写相关反应，则给出：

$$\frac{CR_Y}{S_Y'} = \frac{b_{(A)}R_X}{b_{(P)}S_X} \qquad [19.8a]$$

将 $b = r\ \sigma_Y/\sigma_X$ 和 $R_X/S_X = h_X^2$ 代入，导致：

$$\frac{CR_Y}{S_Y'} = \frac{r_A}{r_P}h_X h_Y \qquad [19.8b]$$

因此，不知道遗传相关 r_A，就不可能使用相关选择差 S_Y'来估计性状 Y 的遗传力，也不能来预测相关反应。等式［19.8］还可写成如下形式：

$$\frac{CR_Y}{S_Y'} = \frac{\mathrm{cov}_{(A)}}{\mathrm{cov}_{(P)}} \qquad [19.8c]$$

这类似于直接反应，$R/S = V_A/V_P$。

间接选择

对相关反应的考察表明选择相关反应比选择期望性状本身有时可能获得更快的进展。换句话说，如果我们想改良性状 X，我们可以选择另一个性状 Y，通过性状 X 的相关反应来获得进展。我们将此称为间接选择；也就是说，选择应用于某一性状而不是期望改良的那个性状。我们将被选择的性状称为次级性状。在什么条件下间接选择将有优势是很容易推导的。如果对期望性状直接选择，那么设 R_X 表示期望性状的直接反应；设 CR_X 表示选择次级性状 Y 时性状 X 的相关反应。于是，间接选择相对于直接选择的优点可用期望反应的比率，CR_X/R_X 表达。从等式［19.5b］中可获得期望的相关反应，从等式［11.4］中可获得期望的直接反应，我们发现：

$$\frac{CR_X}{R_X} = \frac{i_Y h_Y r_A\ \sigma_{AX}}{i_X h_X\ \sigma_{AX}} = \frac{i_Y r_A h_Y}{i_X h_X} \qquad [19.9]$$

从这个表达式可以看出，如果 $r_A h_Y$ 大于 h_X，则间接选择将优于直接选择。这两个数值是两种选择方法的准确度，正如第 13 章所解释的，$r_A h_Y$ 是期望性状 X 的育种值和选择性状 Y 的表型值之间的相关，而 h 是个体选择（即直接选择）的准确度，因此不能期望间接选择优于直接选择，除非次级性状确实具有较高的遗传力并且遗传相关也高。可是，存在可能使间接选择更好的实际情况。可以指出 3 种这样的实际情况。

（1）如果期望性状难于准确度量，因此度量误差降低遗传力，以致于间接选择变得有利。

（2）如果期望性状只能在一种性别中度量，但次级性状可在两性中度量，那么用间接选择可有较高的选择强度。其他条件相同时，间接选择的选择强度将是直接选择的两倍大；但更好的方案将是对一种性别直接选择期望性状，而对另一种性别间接选择次级性状。

（3）期望性状的度量可能很昂贵，例如饲料转化效率。于是选择一个易于度量的相关性状，例如生长速度，则经济上更有利。

对间接选择的详细评估，见 Searle（1965）。下面是间接选择的一个实例，对仅在一种性别中度量的性状，它比直接选择要产生较好的反应。

例 19.2

（数据来自 Nagai 等，1978）对小鼠的两个性状进行选择，度量 12 日龄窝重作为雌性的育仔力，以及个体的 6 周龄体重。育仔力用 N 表示，6 周龄体重用 W 表示。选择在来源不同的两个群体 P 和 Q 中进行，持续了 12 个世代。在每一个群体中，一个系向上选择 N，产生 N 的直接反应和 W 的相关反应；另一个系选择 W，产生 W 的直接反应和 N 的相关反应。反应列于表 19-3，相关反应以斜体表示；所有反应均以 g/世代为单位。

表 19-3　间接选择的反应：以育仔力和 6 周龄体重为例

群　体	P		Q	
选择的性状	N	W	N	W
N 的反应	0.080	*0.134*	0.054	*0.125*
W 的反应	*0.197*	0.680	*0.198*	0.868
观察的 CR_N/R_N	1.675		2.315	
现实的 h^2	0.16	0.40	0.11	0.43
现实的 r_A	0.70		0.73	
$h_W r_A/h_N$	1.11		1.44	
期望的[+] CR_N/R_N	2.2		2.9	

[+]假设 $i_W/i_N=2$。

P 群体的遗传相关，根据等式［19.7］计算如下：

$$r_A^2=\frac{CR_N}{R_N}\frac{CR_W}{R_W}=\frac{0.134}{0.080}\times\frac{0.197}{0.680}=0.485;\quad r_A=0.70$$

Q 群体类似可得 $r_A=0.73$。这样，两个群体之间的 r_A 估值有很好的一致性。

从表 19-3 中 CR/R 的比率可以看出，对育仔力的改良间接选择确实优于直接选择。其理由是 W 比 N 的遗传力高并且 W 可在两种性别中选择。从直接反应中估计的遗传力，作为现实遗传力，列于表中。而且，两个群体对两个性状显示了良好的一致性。实际采用的选择强度未给出，但从选择的比例中可以推导出期望强度。对两个性状选择了相同的比例，但仅对雌性选择了育仔力，而对两个性别选择了体重。因此选择的净强度，体重的预期是育仔力的两倍。

现在用观察的遗传力和遗传相关以及假定的选择强度，根据等式［19.9］能计算育仔力相关反应与直接反应的期望比率。对群体 P，它是：

$$\frac{CR_N}{R_N} = r_A \frac{i_W h_W}{i_N h_N} = 0.70 \times 2 \times \sqrt{\frac{0.40}{0.16}} = 2.2$$

群体 Q 的比率是 2.9。两种情形下，已知的比率有点低于期望值，可能因为对体重已实现的选择强度达不到育仔力的两倍那么大。

虽然上面介绍了间接选择可作为直接选择的一种备择方法，但理论上最有效的方法既不是这种也不是那一种，而是两种方法的结合。将相关性状和期望性状联合起来，作为个体育种值的附加信息使用最为有效。可是，这是将在本章最后一节涉及的更一般问题的一个特殊情形。首先我们将看看间接选择的思想如何扩展到不同环境中的选择。

基因型—环境互作

遗传相关的概念可用来解决与基因型—环境互作有关的一些问题。在第 8 章基因型和环境互作的含义已作过解释，是作为表型值变异的一种来源讨论的，在大多数分析中它与环境方差分不开。这种互作引起的主要问题，以及我们现在讨论的问题，都涉及对当地条件的适应性。基因型—环境互作的存在可意味着在一个环境中最好的基因型在另一个环境中不是最好的。例如，很显然，温带气候下具有最高产奶量的牛种，在热带气候下不一定也具有最高产量。但是环境条件的差异较小是否也需要适应当地的品种，便不那么明显了；同样，如果一个品种被迁移到另一个环境中，那么在原来那种环境中获得的改良会有多少将被转移，在直觉上这也是不明显的。这些问题对育种策略有重要的影响。如果选择是在最好的农场和实验站以及良好的饲养和管理条件下进行的，那么当以后世代被迁移到较差条件中时，所获得的改良会转移过去吗？或者在大多数家畜生活所需要的较差条件下进行选择会更好些？在下面遗传相关的思想将为这些问题的解决提供基础。

在两种不同环境中度量的一个性状，不被看作是一个性状而是作为两个性状。生理机制在某种程度上是不同的，因而为良好表现所需要的基因在某种程度上也是不同的。例如，低营养水平上的生长速度可能主要是饲料利用率的问题，而高营养水平上它可能主要是食欲的问题。将不同环境中的表现视为它们之间有遗传相关的不同性状，依据所掌握的不同性状的遗传力和它们之间的遗传相关，原则上我们能够解决上面提出的问题。如果遗传相关高，那么两种不同环境中的表现几乎表示为同一性状，几乎是由同一组基因决定的。如果遗传相关低，那么性状在很大程度上表现不同，良好的表现需要一组不同的基因。这里我们只考察两种环境，但这一思想可扩展到不定数的不同环境中（Robertson，1959*b*；Dickerson，1962；Yamada，1962）。

让我们考察性状改良从一种环境“转移”到另一种环境的问题。假定我们选择性状 X——比如说高营养水平上的生长速度，要寻求的是 Y 性状——比如说低营养水平上的生长速度的改良。性状 Y 的改良可简单地视为一种相关反应，改良的预期速度可由等式 [19.6] 给出：

$$CR_Y = ih_X h_Y r_A \sigma_{PY}$$

因此，依据每一个环境中表现的遗传力以及这两种表现之间的遗传相关，我们能够预测不同于选择实施环境的另一种环境中性状表现的改良。我们也可以对这种方法预测的改良与我们直接选择性状 Y 预期的改良，即在理想环境中改良的表现，进行比较。这便是前一部分所解释的间接与直接选择的比较。这一比较可由等式 [19.9] 给出的两种预期反应的比率形成，即：

$$\frac{CR_Y}{R_Y} = r_A \frac{i_X h_X}{i_Y h_Y}$$

这一比率指明了在不是被改良群体所需的某种环境中实施选择，我们可有多少改良的增益或损失。如果我们假定选择强度不受实施选择的环境影响，那么如果 $r_A h_X$ 大于 h_Y，则间接方法会更好，其中 h_X 是选择环境中遗传力的平方根，h_Y 是群体后来生活所需环境中遗传力的平方根。如果遗传相关高，那么这两个性状大体上可视为同一性状；如果没有影响遗传力和选择强度的特殊环境，那么在哪一种环境中进行选择都几乎不会产生什么差异。但如果遗传相关低，那么在群体注定要生活的那种环境中进行选择会有利，除非在另一种环境中遗传力和选择强度相当高。

这是处理不同环境中选择的理论基础。已有几个实验在验证这一理论。通常，发现相关反应小于直接反应，即如果选择在改良所要求的环境中进行则选择最有效时，他们便证实了这一理论。可是，这些实验没有摆脱以前提到的不一致，特别是上选和下选反应间的不一致。例 19.3 简要描述了小鼠的一个实验；其他例子见赤拟谷盗（*Tribolium*）的两个实验，性状是幼虫生长（Yamada 和 Bell，1969）和产卵率（Orozco，1976）。

环境敏感性　在第 8 章对环境敏感性差异引起的基因型—环境互作方式作过解释。为预测选择反应之目的，遗传相关提供了互作数量化的方法。但是，考虑环境敏感性可以帮助我们理解不同环境下的选择反应。高遗传相关意味着所有基因型几乎以同样的方式对环境差异作出反应；像图 8-2 那样有几乎完全平行的回归直线。低遗传相关意味着基因型反应不同，具有不同斜率的回归直线，即个体有不同的环境敏感性。选择怎样对这些敏感性差异起作用呢？根据它们是促进还是抑制这个性状，将环境归属于“好”或“坏”，这是很方便的；实践中通常是追求增加，因此增加性状的环境谓之为“好”。从图 8-2 中可以看出选择对敏感性的影响。好环境中的向上选择倾向于选择高敏感性的个体，坏环境中的向下选择具有同样的效果。相反，坏环境中的向上选择和好环境中的向下选择倾向于选择低敏感性的个体。两种类型的选择可以区分，当选择和环境同在一个方向上起作用（好环境中上选或坏环境中下选）时，为协同选择（synergistic selection）；相反，当它们在相反方向上起作用时，为颉颃选择（antagonistic selection）。因此，敏感性将被协同选择增加，被颉颃选择降低。这些期望结果已被 Falconer（1990）综述的许多实验所证实。

例 19.3

数据来自 Falconer（1960*b*）。根据两种日粮对小鼠 3～6 周龄的生长进行选择："好"日粮即是正常日粮，"坏"日粮是正常日粮掺杂进 50% 不能消化的纤维。实验开始时，坏日粮降低生长大约 20%。每一世代都度量直接和相关反应，直接反应来自在选择日粮中生长的第一窝后代，相关反应来自在另一种日粮中生长的第二窝后代。选择以两个方向进行。相反方向选择之间以及早期和晚期世代之间存在不一致性。为例证理论的目的，仅取前 4 代的结果和上选与下选之间的歧化表示反应可避免这些不一致。表 19-4 给出了根据等式［19.9］由每对系群单独以及根据等式［19.7］由两对系群一起计算遗传相关所需的信息。反应是 g/世代。正如所预料的那样，两个相关反应都低于直接反应，两者之间的比率表明遗传相关为 0.66。

表 19-4 到第 4 世代的每代歧化（g）

	性 状	
	好日粮下生长	坏日粮下生长
选择强度，i	1.66	1.40
现实遗传力，h^2	0.41	0.36
直接反应，R	0.90	1.20
相关反应，CR	0.48	0.98
由等式［19.9］得到的遗传相关，r_A	0.67	0.65
由等式［19.7］得到的遗传相关，r_A	0.66	

实践中，我们想要的经常不是特定环境中的表现，而是一定环境范围中的表现，包括好环境和坏环境，即不同环境中的良好平均表现。经常不能在一种以上的环境中度量个体。因此，如果要采用个体选择，便出现是通过协同选择来选择高敏感性还是通过颉颃选择来选择低敏感性的问题。Falconer（1990）的综述表明通过颉颃选择可获得最佳的整体表现，这种情况更经常发生（21 种情形中有 14 种）。可是，理论的检查发现没有理由预料会发生这样的结果。如果能够采用家系选择，则家系可在环境之间分化，可以估计出每个家系的平均表现。为获得平均表现的最大改良，James（1961）已指明如何用一个指数将一个家系的两种表型联合的细节。

例 19.4

例 19.3 的小鼠实验提供了有关环境敏感性和总体表现的数据。有第 7 世代，度量了一个未选择的对照，并在表 19-5 中以与对照离差的形式给出了两个方向上的反应。相关反应以斜体字表示。每一个系的环境敏感性是两种日粮下小鼠生长之间的差异，表示在"日粮效应"下面。正如所期望的，最敏感的系群是那些在好日粮中上选以及在坏日粮中下选，即协同选择的系群。像大多数实验一样，在"两种日粮平均"下面表示的、具有最佳的整体或平均表现的系群是在坏日粮中向上选择的以及在好日粮中向下选择的，即颉颃选择的系群。

表 19-5　以与对照离差形式表示的第 7 世代的总反应（g）

选　择		反　应			敏感性
方向	日粮	好日粮中生长	坏日粮中生长	两种日粮平均	日粮效应
上	好	2.3	*0.6*	1.45	5.4
上	坏	*1.6*	3.1	2.35	3.5
下	好	-2.8	*-2.9*	-2.85	3.6
下	坏	*-1.2*	-3.2	-2.20	6.8

诸如日粮等某种已知因素不同的环境被称为大环境（macro-environments）。可是，一单个大环境内部有小环境差异，有关选择如何影响敏感性的结论也适用于一单个大环境内的小环境方差。根据表型选择的个体具有良好的基因型也有过良好环境的经历。因此选择是协同的，如果存在基因型—环境互作，那么在两个方向上的选择都将增加环境敏感性。增加的敏感性表现为环境方差 V_E 的增加以及遗传力的减少。在很多选择实验中看到的表型方差增加可能部分地归于这种原因，正如在第 12 章中已指出的那样。

指数选择

当选择应用于动植物经济价值的改良时，一般是同时应用于几个性状，而不仅仅是一个性状，因为经济价值依赖于一个以上的性状。这通常被称为*多性状选择*。例如，从一个猪群中获得的利润依赖于它们的生殖力、哺育能力、生长速度、饲料利用效率和胴体品质。那么，为了获得经济价值的最大改良，选择应当怎样应用于组分性状呢？有几种可能的方法。一种方法是在连续世代中单独地对每一个性状依次进行选择（*顺序选择*）；或者一种方法是在同一时间对所有性状进行选择，但是独立地，剔除没有达到每一个性状的某一特定标准的所有个体，而不管这些个体其他任何性状的值（*独立淘汰水平*）。然而，预测指明产生经济价值最快改良的方法是同时一起对所有的组分性状进行选择，并按其相对经济重要性、遗传力以及不同性状之间的遗传和表型相关给每一个性状以适当的权重。因此对经济价值选择的实践存在一些复杂问题。组分性状必须在一起合并成一个得分或指数，以这样一种方式应用指数选择，好像指数是一个单个性状一样，将产生经济价值可能最快的改良。

指数选择的原理在第 13 章介绍过，这里就不完全重复了。这里需要指出的指数中的主要差异是，要预测的不是一单个性状的育种值，而是以经济方式评估的几个性状的复合育种值，因此比第 13 章中制订的指数更复杂。可是，我们开始将考虑运用指数改良一单个性状的较简单问题，然后将这种思路扩展到提高经济价值。实践中，正如第 13 章所解释的那样，BLUP 可用来估计实际数据组中的固定效应和解释其他复杂情况，同时评估价值的育种值。

指数的构建

选择的目标，无论它是什么，都将被称为*价值*（merit），价值的育种值将用符号 H 来表示。与以往一样，为改良价值，其指数构建是：

$$I = b_1P_1 + b_2P_2 + \cdots + b_mP_m \qquad [19.10]$$

其中，P_1 到 P_m 是 m 个性状的表型度量值，它们是选择的基础；b_1 到 b_m 是待定的相应加权因子。b 是 H 对 I 的偏回归系数。指数中包括来自亲属的信息，因此，P 能是第 13 章中解释的方式中的亲属度量值。

单性状 首先考察旨在只改良一个性状的选择。于是应用指数选择的目的是用次级性状作为改良这个期望性状的辅助。指数方程，它的解给出指数中 b 的值，与等式 [13.10] 完全一样，其中性状 1 作为要改良的性状。

$$\left.\begin{aligned} b_1P_{11} + b_2P_{12} + \cdots + b_mP_{1m} &= A_{11} \\ b_1P_{21} + b_2P_{22} + \cdots + b_mP_{2m} &= A_{21} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ b_1P_{m1} + b_2P_{m2} + \cdots + b_mP_{mm} &= A_{m1} \end{aligned}\right\} \qquad [19.11]$$

这里的符号与第 13 章中的一样缩写了。例如，P_{11}是性状 1 的表型方差，P_{12}是性状 1 和 2 的表型协方差；类似地，A_{11}和 A_{12}是加性遗传方差和协方差。方差和协方差可以用遗传力和相关的形式表示如下，其中下标 i 和 j 指任何两个不同的性状，σ^2 是表型方差：

$$\left.\begin{aligned} P_{ii} &= \sigma_i^2 & A_{ii} &= h_i^2\,\sigma_i^2 \\ P_{ij} &= r_P\,\sigma_i\,\sigma_j & A_{ij} &= r_A h_i h_j\,\sigma_i\,\sigma_j \end{aligned}\right\} \qquad [19.12]$$

当代入方差和协方差的值时，等式 [19.11] 的解提供了在等式 [19.10] 指数中要使用的加权因子 b 的数值。指数的构建随后示例在例 19.5 中。在解释不同形式的指数方程之后，我们将涉及选择的期望反应。

经济价值 下一步考察经济价值的提高。经济价值来自个体销售所获得的利润。在实际的育种操作中，经常可能对个体指定经济价值。于是，这是价值（merit）的表型值，便是要改良性状的表型值，并且指数正是为这单个性状的改良而构建。但是，指数方程，其解作为指数中 b 值的指数方程，在一个方面与上述不同。在考虑对个体选择的同时，它们的经济价值是不知道的，因此不能作为一个性状包括在指数中。于是指数方程如下。为了便利比较，在这种情况的价值中，仍然视性状 1 是要改良的性状。

$$\left.\begin{aligned} b_2P_{22} + b_3P_{23} + \cdots + b_mP_{2m} &= A_{21} \\ b_2P_{32} + b_3P_{33} + \cdots + b_mP_{3m} &= A_{31} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ b_2P_{m2} + b_3P_{m3} + \cdots + b_mP_{mm} &= A_{m1} \end{aligned}\right\} \qquad [19.13]$$

方差和协方差必须从经济价值过去的记录和指数中性状值中估计。

多性状 最后，考察对几个性状的同时选择。目标是提高综合育种值（aggregate breeding value）或净价值（net merit），它是要改良的所有性状的一个特定组合。价值现在被定义为：

$$H = a_1A_1 + a_2A_2 + \cdots + a_nA_n \qquad [19.14]$$

这里 A 是要改良的 n 个性状的育种值，a 是加权因子，它表示育种者加给每一个性状的相对重要性。加权因子可以是经济价值；也就是说，每一个 a 是 1 个单位性状的货币单位数值。当没有个体经济价值的记录，以致于上面描述的指数不能使用时，如果目的是提高经济价值，那么该如何构建指数。但是，给性状指定货币值不一定是改良的最佳方法。

在有关猪的改良方面，Fowler，Bichard 和 Pease（1976）讨论了加权的其他标准。如果加权因子不是用货币单位表示的，它们必须以某种其他方式表示每个性状增加 1 个单位对育种者的相对重要性。Yamada，Yokouchi 和 Nishida（1975）描述了以这种方法构建的指数。

价值定义中（等式［19.14］）和指数中（等式［19.10］）的性状数目可以不同：可能存在一些性状，它们不在 H 中，但如果包括在 I 中，则通过它们的相关而能有助于提高 H；相反，它们可以是不能度量的 H 中的性状，因此也不在 I 中。可是，重要的是要注意到，如果目的是要提高经济价值，那么影响经济价值的所有性状必须包括在 H 的定义中。

用第 13 章描述的同样方式，令价值与指数间相关 r_{HI}最大化，可以获得其解为指数中要用的 b 值的指数方程。它们如下：

$$\left.\begin{aligned} b_1P_{11}+b_2P_{12}+\cdots+b_mP_{1m}&=a_1A_{11}+a_2A_{12}+\cdots+a_nA_{1n}\\ b_1P_{21}+b_2P_{22}+\cdots+b_mP_{2m}&=a_1A_{21}+a_2A_{22}+\cdots+a_nA_{2n}\\ &\cdots\cdots\\ b_1P_{m1}+b_2P_{m2}+\cdots+b_mP_{mm}&=a_1A_{m1}+a_2A_{m2}+\cdots+a_nA_{mn}\end{aligned}\right\} \quad [19.15]$$

借助等式［19.12］，方差和协方差可再次用遗传力和相关的形式表示。下面的例 19.6 示例了从这些等式中构建一个指数，通过仅考察两个性状而简化了。

实践　前面对指数的解释用于阐明构建指数的原理。实践中，复杂困难的计算可由计算机来进行。REML 程序估计了所需要的群体参数，BLUP 计算了被选择个体间的个体综合育种值。在第 13 章中已概述了这种方法。

反　应

指数方程（［19.11］，［19.13］，［19.15］）是用这样一种方法被度量的，即价值对指数值的回归是 1，即 $b_{HI}=1$。因此指数（等式［19.10］）中的 b 值要校正，以便指数 I 中的度量值在数值上符合表示价值 H 的单位，当两者都是偏离平均值的离差时，无论它们是什么。用这种方法，指数变成价值育种值的一种预测值。由于 $b_{HI}=1$，它表明 $r_{HI}=\sigma_I/\sigma_H$，就像第 13 章中显示的那样。选择的预期反应与第 13 章中的相同，价值的预测变化由下式给出：

$$R_H=ir_{HI}\ \sigma_H \quad [19.16]$$

或者，如果指数没有重新度量，则有：

$$R_H=i\ \sigma_I \quad [19.17]$$

当价值中只有一个性状时，即单性状选择时，指数的方差与等式［13.12］的相同，从中可估计 σ_I。扩展到包括多性状选择，则指数方差可表示如下。为了简化记号，让 $\sum_1 aA$ 表示［19.15］中第一个方程右边各项的总和，$\sum_2 aA$ 表示第二个方程右边各项的总和，等等。于是方差为：

$$\sigma_I^2=b_1\sum_1 aA+b_2\sum_2 aA+\cdots+b_m\sum_m aA \quad [19.18]$$

指数的标准差 σ_I 为比较提高价值的不同指数的相对效率提供了一种简单的方法，因为，正如从等式［19.17］中可看到的，价值的反应简单地与 σ_I 成比例。

指数或价值的任何一个组分性状的反应可预测如下。假定我们想要预测性状 1 的反应。这是性状 1 对指数选择的相关反应，它可通过改写等式［19.5*a*］来预测，把性状 1 放在 Y 的位置上，指数 *I* 放在 X 的位置上。指数的反应与等式［19.17］给出的价值反应是相同的。于是，性状 1 的相关反应变为：

$$CR_1 = b_{(A)1I} i\ \sigma_I = \frac{\mathrm{cov}_{(A)1I}}{\sigma_I^2} i\ \sigma_I = \frac{i}{\sigma_I}\mathrm{cov}_{(A)1I} \qquad [19.19]$$

这里 $\mathrm{cov}_{(A)1I}$是性状 1 和指数的加性遗传协方差，它的获得如下。将等式［19.10］乘以 A_1，给出 A_1 和 I 的乘积和，为 $b_1P_1A_1 + b_2P_2A_1 + \cdots$，现在，如果每一个 P 写成（$A + E$），剔出乘积 AE，因为育种值和环境是不相关的。现在可以看出所需要的协方差如下，其中方差和协方差以指数方程的符号书写：

$$\mathrm{cov}_{(A)1I} = b_1A_{11} + b_2A_{12} + \cdots + b_mA_{1m} \qquad [19.20]$$

对指数的构建这些方差和协方差一定是已知的，因此代入等式［19.19］可给出预测反应。

下面两个例子将使在指数构建中包含什么更为清楚，并将引出上面没有解释但却很有意义的一点或两点启示。

例 19.5

用例 19.6 描述的实验来示例改良单个性状的指数运用。假定要改良的性状是小鼠的体重，我们考虑在指数中使用相关性状尾长。让体重作为性状 1，尾长作为性状 2。构建指数所需要的参数给在表 19-6 中，这些值取自 Rutledge，Eisen 和 Legates（1973）。

表 19-6 小鼠体重与尾长的有关参数

性状	h^2	h	σ_P^2	σ_P	r_A	r_P
1 = 体重(g)	0.36	0.60	6.37	2.52	0.29	0.45
2 = 尾长(cm)	0.44	0.67	0.28	0.53		
$P_{11} = 6.37$	$P_{22} = 0.28$		$P_{12} = P_{21} = 0.6010$			
$A_{11} = 2.2932$	$A_{22} = 0.1232$		$A_{12} = A_{21} = 0.1557$			

根据等式［19.11］，求解的指数方程是：

$$b_1P_{11} + b_2P_{12} = A_{11}$$
$$b_1P_{21} + b_2P_{22} = A_{21}$$

根据等式［19.12］计算的方差和协方差的值给出在表中。将这些值代入上面的方程中给出：

$$6.37b_1 + 0.6010b_2 = 2.2932$$
$$0.6010b_1 + 0.28b_2 = 0.1557$$

解为：

$$b_1 = 0.386;\ b_2 = -0.272$$

因此，选择指数为：

$$I = 0.386W - 0.272T$$

其中，W 和 T 分别为体重（g）和尾长（cm）。为方便起见，除以 b_1，可重新度量指数，这给出：

$$I' = W - 0.705T$$

指数值重新度量后有所改变，但个体价值的次序没有改变。

注意到尾长在增加体重的指数中给予了负的权重。换句话说，尾长是体重的环境指标而不是育种值的指标。其理由是环境相关为 0.56，比遗传相关 0.29 高得多。这例证了以前没有提及的一点，即在指数中一个性状可用作环境离差指标而不是育种值的指标。

次级性状的有用性可从它的权重系数 b_2 来判断。如果性状 2 对性状 1 的遗传和表型回归是相同的；或者就相关而言，如果 $r_A/r_P = h_1/h_2$，则可表明 $b_2 = 0$。在这些条件下，一个次级性状将不产生益处；实际上，估计的误差将使它比无用性更糟（Sales 和 Hill，1976）。

为了预测选择反应，我们必须计算指数的方差。为此目的，必须使用未重新度量的指数。根据等式［19.18］，方差是：

$$\sigma_I^2 = (0.386 \times 2.2932) + (-0.272 \times 0.1557) = 0.8428$$

并且，$\sigma_I = 0.918$。于是，如果选择强度已知，根据等式［19.17］就可以预测反应。为了了解次级性状如何有用，假定选择强度相同，我们可以比较指数选择和简单选择的期望反应。根据等式［11.3］，简单选择的反应是 $ih^2\sigma_P$。因此，指数选择/简单选择反应的比率是 0.918/0.907 = 1.012。指数选择仅比只对体重的简单选择好 1%。

最后，由指数选择引起的尾长的预期变化将是多大？性状 1 的相关反应由等式［19.19］和［19.20］给出，但这种情形下，我们想要的是性状 2 的反应。首先，根据等式［19.20］，

$$\mathrm{cov}_{(A)2I} = b_2A_{22} + b_1A_{21} = -0.0335 + 0.0601 = +0.0266$$

然后，根据等式［19.19］，尾长的预期反应是：

$$CR_2 = +i(0.0266/0.9185) = +0.03i\text{cm/世代}。$$

预期尾长有一个很小的增加。开始时似乎指数中尾长的负权重应当导致下降。但相关反应取决于遗传相关，而它是正的。

例 19.6

在例 19.5 取得数据的小鼠实验中以改变体重和尾长为目标应用了指数选择并比较了观察和期望反应（Rutledge，Eisen 和 Legates，1973）。此例的目标是通过增加一个性状和减少另一个性状来改变体型结构。对 4 个系选择了 7 个世代，两个系选择增加体重和降低尾长，两个系在相反的方向上选择。我们将解释增加体重和降低尾长的指数构建。需要的参数在例 19.5 的表中给出了。另外，我们需要两个性状的经济权重。给每一个性状变化的一个标准差指定相等的经济价值。因此，指定的权重 a 是表型标准差的倒数，它们是 $a_1 = 0.40$，$a_2 = -1.89$。

根据等式［19.15］，指数是：

$$b_1P_{11} + b_2P_{12} = a_1A_{11} + a_2A_{12}$$
$$b_1P_{21} + b_2P_{22} = a_1A_{21} + a_2A_{22}$$

代入方差和协方差以及权重，导致：

$$6.37b_1 + 0.601b_2 = 0.9173 - 0.2943 = 0.6230$$

$$0.601b_1 + 0.28b_2 = 0.0623 - 0.2328 = -0.1706$$

其解是：

$$b_1 = 0.195; b_2 = -1.027$$

因此选择指数为：

$$I = 0.195W - 1.027T$$

其中，W 和 T 分别是个体的体重和尾长。对相反方向的选择，即降低 W 和增加 T，指数中的符号可简单地反过来。

根据等式[19.18]，指数的方差是：

$$\sigma_I^2 = (0.195 \times 0.6230) + (-1.027 \times -0.1706)$$
$$= 0.2967$$

$$\sigma_I = 0.5447$$

对4个系求平均，现实的选择强度是 $i = 1.01$。根据等式[19.17]，以实际应用的选择为基础的预期反应是：

$$R_H = 1.01 \times 0.5447 = 0.55 \text{ 指数单位/世代。}$$

在选择增加 W 和降低 T 的系中观察的反应是0.26和0.30，在相反方向选择的系中则为0.42和0.45。观察反应有点低于预期反应的理由可能是构建指数的参数估计不准确。

（本实验中计算出的一些数据与原始论文中给出的那些数据有些不同。其理由是论文中使用的参数是从选择前的基础群体中获得的，但这里采用的公布参数是用选择共同保持的未选择系的参数。）

选择对遗传相关的影响

在我们结束本章之前，需要讨论多性状选择的一个重要后果。就像经一段时间选择之后遗传力预计要变化一样，遗传相关也是这样。如果同时对两个性状选择，那么由于下面的原因，预料最终它们之间的遗传相关会变成负的。在有利方向上影响两个性状的那些多效性基因因选择将强烈地起作用，并很快地趋向固定。于是它们对两个性状的方差或协方差贡献甚微。可是，对一个性状有有利影响而对另一个性状有不利影响的多效性基因并不那么强烈地受选择的影响，将较长时间保持在中间频率状态。因此，两个性状剩余协方差的大部分是由这些基因引起的，由此产生的遗传相关将是负的。负遗传相关，不管它是用这种方法由选择产生的还是从一开始就有的，其后果是两个性状的每一个都可能表现出远离0的遗传力，尽管同时对它们选择时两者都没有反应。在第12章我们已经讨论过了，人工和自然选择的联合效应的结果基本是相同的状态：选择极限已达到，即使对性状实施人工选择仍具有巨大数量的加性遗传方差。

在几个实验中已检验过选择将改变遗传相关的理论期望，但正像遗传相关经常变化一样，证据却是互相矛盾的。这个问题的综述和讨论参见 Sheridan 和 Barker(1974)。

习　题

19.1　下面的数据取自一群肉仔鸡的同胞分析。它们涉及单个雄性从 5 到 9 周龄的增重(G)和同一时期消耗的饲料(F)重量,两者都以克为单位。给出的数字是半同胞家系之间的方差和协方差组分,以及总表型方差和协方差。从这些数据中,计算两个性状的遗传力,以及它们之间的表型、遗传和环境相关。

	方　差		G 和 F 的协方差
	增重	饲料消耗	
公畜间	1 602	6 150	2 229
总的	12 321	61 504	22 848

数据基于 Pym, R. A. E. & Nicholls, P. J. (1979) *Brit. Poult. Sci.*, **20**, 73 ~ 86。

[解答序号:40]

19.2　数据涉及果蝇(*Drosophila*)的两个性状:身体大小(度量为胸长)和繁殖力(度量为 4 天内产卵数)。在一个遗传上变化的群体中和在一个由近交系之间杂交的 F_1 个体组成的遗传上一致的群体中度量了表型方差和协方差。从这些数据中,估计可变化群体中的 3 种相关——表型、遗传和环境相关。这里遗传相关的含义如何不同于习题 19.1 中的遗传相关的含义?

群体	方　差		协方差
	身体大小	繁殖力	
变化	0.366	43.4	0.87
一致	0 186	16.6	0.27

数据来自 Robertson, F. W. (1957) *J. Genet.*, **55**, 428 ~ 443。

[解答序号:50]

19.3　再次考虑习题 19.1 中有关肉仔鸡的数据。假定在一个系中应用选择来增加增重,在另一个系中增加饲料消耗。两个系中选择的比例为雄性 10%,雌性 20%。当直接选择时和当作为相关性状起反应时,计算两个性状每世代的预测反应。

[解答序号:60]

19.4　对习题 19.1 描述的肉仔鸡群进行了 5 个世代的选择。一个系选择增加增重(G),另一个系选择增加饲料消耗(F)。5 个世代采用的总选择差和总反应如下,反应是偏离一对照系的离差。

	选择的系	
	G	F
选择差(g)	574	1 312
G 的反应(g)	186	120
F 的反应(g)	412	525

计算两个性状现实遗传力以及它们之间的现实遗传相关。

数据来自 Pym, R.A.E.& Nicholls, P.J.(1979) *Brit. Poult. Sci.*, **20**, 73 ~ 86。

[解答序号:70]

19.5 通过对雌性窝产仔数的选择，或者通过对两个性别的体重进行选择可以增加小鼠的窝产仔数。给定下面的参数，这两种简单方法哪一种更好?

窝产仔数的遗传力	= 0.22
体重的遗传力	= 0.35
遗传相关	= 0.43
选择比例:	雌性 = 25%
	雄性 = 10%

[解答序号:80]

19.6 除了习题 19.5 中考虑的增加窝产仔数的两种方法之外，第三种方法是选择雌性的窝产仔数和选择雄性的体重。有几种方法可用来实现它。假定方法如下。雄性在适当的年龄(例如 6 周龄)称重并从大数目中选择 25%。中选雄性的每一个都与窝产仔数未知的随机选择的一组 4 个年轻雌性进行交配。当雌性产生它们第一窝仔畜时，在与每一雄性交配的 4 个中选择最好的一个雌性。这些中选的雄性和雌性的窝仔畜被饲养作为下一世代。这种方法比仅选择雌性窝产仔数要好多少? 采用习题 19.5 中给定的遗传力和遗传相关。

[解答序号:90]

19.7 数据涉及习题 19.1 中描述的肉仔鸡。假定使用饲料消耗重量(F)作为选择的辅助，期望提高 5 ~ 9 周龄的生长(G)。计算评估被选择亲本的适宜指数。所需参数的估计值如下。为了带小数点以便于计算，这里重量的单位从克改变为 100 克单位。

	G		F
h^2	0.52		0.40
σ_P	1.11		2.48
r_A		0.71	
r_P		0.83	

[解答序号:100]

19.8 如果应用习题 19.7 中计算的指数来选择生长，假定选择强度和习题 19.3 中的一样为 $i = 1.5775$，那么每世代预测的改良会是多少? 对提高生长来说，指数选择比没有一个次级性状的辅助仅选择生长要好多少?

[解答序号:110]

19.9 如果像习题 19.8 中那样，采用指数选择生长，则次级性状饲料消耗变化的预期速度将是多大?

[解答序号:120]

19.10 从习题 19.7 的数据和计算中，计算提高经济价值的指数，假定生长的值是每 100g 增重为 8 分，饲料消耗的值是每 100g 为 - 2 分。

数据来自 Pym, R.A.E.& James, J.W.(1979) *Brit. Poult. Sci.*, **20**, 99 ~ 107。

[解答序号:130]

19.11 像习题 19.3 那样，当 10% 的雄性和 20% 的雌性被选择时，计算从采用习题 19.10 的指数选择中预期的经济价值改良的速度。对改良经济价值，指数选择比仅选择生长要好多少?

[解答序号:140]

第 20 章　自然选择下的数量性状

关于数量性状遗传特性的方方面面讨论已占了本书的主要部分，但我们极少注意到自然选择的影响，现在必须对此做一些说明以补救这种忽略。数量遗传学这门学科理论发展的一个条件是，指定存活力和生殖力没有差异，也就是说，假定自然选择不存在。虽然对很多目的这种假设可能不会导致严重的误差，但直到计入自然选择的影响，我们才能对数量性状有完整的理解。但是，与可以使简单的理论描绘以及以此为基础的预测发生困惑的复杂性相比，自然选择对数量性状的作用有更加广泛的意义。对迄今为止我们没加说明就已接受的数量性状遗传特性必须找到一种解释，这便是自然选择。一个群体的遗传特性是过去自然选择、突变和随机漂变的产物。我们必须用这个过程对存在的遗传变异做出解释，而且必须主要用自然选择对性状的遗传特性差异，例如，与其他性状相比，某些性状的加性方差呈比例地增多；一些性状表现近交衰退，而另一些性状则没有等事实做出说明。然而这是很广泛的问题，对它们的解答仍然很遥远，本章只能对其本性略加说明。在考察自然选择影响数量性状的方式之前，我们将说明自然选择本身及其含义是什么。

自然选择

适合度及其组分

自然选择所选择的“性状”是适合度。一个个体的适合度是它传给下一世代的基因贡献，或者是出现在下一世代中的后裔数目。相对适合度是一个个体相对于群体平均值的适合度，即 $W/\overline{W}$，假如 W 是个体的适合度。如果一个群体其数目既不增加也不减少，那么它的个体平均适合度是 1，于是绝对适合度和相对适合度相同。精确地定义适合度有困难，这在于将个体的适合度与其父母的适合度区别开来。例如，在哺乳动物中，幼年后裔的存活部分地取决于自身的存活力，这是它们自己适合度的一个方面，也部分地取决于它们接受到的双亲照顾，而这是双亲适合度的一个方面。适合度在世代间的这种交错意味着在个体的生命中没有一个精确的时间使我们可以说它的特性反映了它自己的适合度而不是它父母的适合度。因此世代间分开的那一点必定是一种或多或少武断的选择。

一个群体的平均适合度是必须十分小心使用的一个概念。上面已经说过，如果群体数目既不增加也不减少，那么该群体中个体的平均适合度为 1。这显得足够简单了。但是一个群体的数目是否增加或减少或保持稳定在很大程度上取决于它可利用的环境资源。一个群体内个体之间的自然选择可能改变群体的遗传组成，但如果群体已经处于其环境容纳量的极限时，则平均适合度将不改变。以下提到平均适合度时，将假定群体不受环境资源的限制。

一个个体的适合度是它的发育和生理所有过程的最终结果。个体之间在这些过程中的差异可从能度量特性的变异中看出，这种特性能作为数量性状研究。因此，每一个数量性

状的变异在或大或小的程度上反映了适合度的变异；适合度的变异在理论上可分解成数量性状的变异。例如，考察一种哺乳动物如小鼠。图 20-1 例释了对雌性适合度有贡献的性状阶层系统。显示的所有性状几乎都作为数量性状在遗传上研究过。适合度本身能分解成两个主要组分，产生后裔的总数和这些后裔的质量，后者可根据它们的断奶重来度量。主要组分的变异，如果得以恰当地度量，将会解释适合度的所有变异。主要组分的变异可依次归因于其他性状，它们中的一些显示在简图的第 3 栏。这些又受其他一些性状的影响，其中有几个显示在第 4 栏。第 4 栏中的性状本身又受很多其他性状的影响。例如，在这些性状之间有各种促性腺激素排出量等生理机能，它们影响排卵率、胚胎存活和泌乳量。另外，有一些性状，它们对适合度的影响不那么直接，也不那么明显，但它们与适合度的一些组分相关联。例如，身体大小可能与第 4 栏中几个或所有性状相关联。我们必须检查的问题是所有这些性状如何受适合度选择的影响，以及自然选择的作用又如何与性状在阶层系统中的位置相关联。

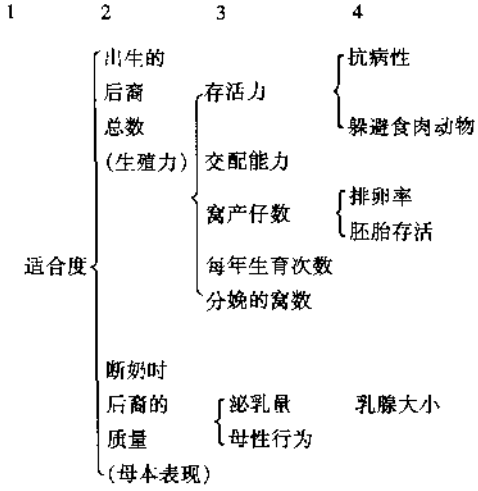


图 20-1 哺乳动物例如小鼠适合度的一些组分，以显示变异原因的阶层系统。这些数量性状的每一个变异或大或小都与适合度的变异有关

适合度的度量 直接度量适合度，尤其是个体的适合度是很难的。分别度量适合度的主要组分不很困难。于是综合的适合度可通过合并组分的值来估计。从人口统计学的研究中可获得这类信息，在人口统计学中一个社区（cohort）从出生到死亡都被跟踪，可获得年龄特异性的生存情况、生殖力、交配能力等数据。在人类的工业化社会中，无疑适合度最重要的组分是养育孩子的数目，因为童年和生殖结束之间的死亡率很低。因此完整的家庭大小为双亲的适合度共同提供了一种相当好的度量，当然，尽管它将婴儿和童年的存活包括在父母的适合度中。通过考虑生殖率（即孩子出生时父母的平均年龄）可以获得一种改进的度量（见 Waller，1971）。将所有组分合并成一单个值的适合度度量，即“竞争指数”，已应用于果蝇（*Drosophila*）中。所关注的品系个体与一个作过遗传标记的测交品系竞争性地饲养在一起，以便能够识别杂交产生的后裔。竞争指数是两个纯系产生的后裔相对数目。这个方法的细节和估计果蝇适合度的其他方法以及参考文献见 Sved（1989）。

数量性状和适合度之间的关系

正如上面描述的，现在我们讨论自然选择影响不同种类性状的方式。我们还需考虑对性状采用的是什么种类的选择，这将对有关基因的频率产生什么影响，从而性状的遗传特性又如何受到影响。

“适合度轮廓”

如果性状表型值不同的个体其适合度存在差异的话，那么自然选择就对这个数量性状起作用。从被称为“适合度轮廓”的图形中可以确定自然选择的类型和它的强度，该图是用适合度对表型值所作。虽然有相当的困难，但从对一个社区由生至死所做的数量性状表型和适合度组分研究中可经验性地获得这种信息（例如，Clutton－Brock，Guinness 和 Albon，1982；Grant，1986）。更普遍的是，相对适合度可从选择前后性状表型频率分布的变化中推断出来。适合度轮廓图解绘制在图20-2中。关于果蝇（*Drosophila*）适合度轮廓

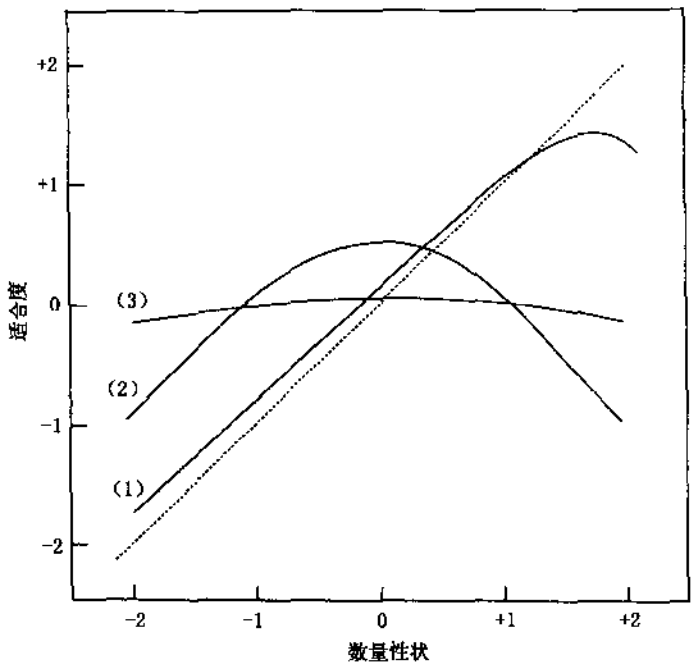


图 20-2　正文中解释的“适合度轮廓”。两个坐标轴上的尺度都是距平均值的标准差

的实例见 Mackay（1985*a*）。如果度量的数量性状是适合度本身，则将获得的轮廓是点状对角线。曲线（1）是在对高性状值强烈定向选择下的轮廓，正如对适合度主要组分所期望的那样。轮廓（2）代表经历稳定的自然选择的性状。存在一种中间最优度，因为最适应的个体是那些性状值为平均值或接近平均值的个体。歧化选择下的性状适合度轮廓没有显示：它是稳定选择轮廓的颠倒，选择不利于中间型。最后，轮廓（3）代表处于很弱选择下的性状，它们很近于中性带有全部观察表型范围的相应适合度。在这类性状具有不同

值的个体之间适合度几乎没有差异。先前几章中，性状被指定与适合度“紧密联系”，但对“紧密联系”的确切意义并没有解释。现在可以看出，与适合度紧密联系的性状是具有像曲线（1）那样适合度轮廓的主要组分。随着轮廓接近曲线（2）的轮廓，联系的紧密程度降低了。

对自然选择的反应

适合度

适合度自身对选择的反应已有极好的解答，被称为*费歇尔基本法则*，它表述为任何时候适合度的增加等于那个时候适合度的加性遗传方差。这个法则引发了关于它的有效性和它的一般性的大量讨论；见 Frank 和 Slatkin（1992）以及 Edwards（1994）。Price（1972）给出了这个法则的一个证明和为什么它引起这么多困难的解释。通过考虑加权选择差，在没有世代重叠的群体中，对一单个世代证实这个结论很简单，结果如下。

数量性状的选择反应可通过等式［11.2］（$R = h^2S$）从选择差中来预测。选择差 S 是中选亲本的加权平均优势，权重是后裔的相对贡献，据此反应得以评估。适合度必须表达为相对适合度。在那个世代中，如果 k 是任一特定个体的后裔数目，$\overline{k}$ 是所有个体后代的平均数，那么个体的相对适合度是 $W = k/\overline{k}$。平均相对适合度为 $\overline{W} = 1$。相对适合度也是给个体对选择差贡献（$W - \overline{W}$）指定的权重。因此，如果 N 是亲本世代中个体的总数，那么相对适合度的加权选择差是：

$$\begin{aligned} S_W &= \frac{\sum W(W - \overline{W})}{N} \\ &= \frac{\sum W^2}{N} - \frac{\sum W}{N}\overline{W} \\ &= \overline{W^2} - (\overline{W})^2 \\ &= V_{P(W)} \end{aligned} \qquad [20.1]$$

（推导的最后一步与等式［3.4］联系在一起解释。）这样适合度的选择差等于适合度的表型方差，我们可能意外地注意到选择强度等于表型标准差：

$$i_W = S_W/\sigma_W = \sigma_W \qquad [20.2]$$

将等式［20.1］代到等式［11.2］中表明适合度对自然选择的反应就等于适合度的加性遗传方差：

$$R_W = h_W^2 V_{P(W)} = V_{A(W)} \qquad [20.3]$$

根据这个法则预测，适合度的任何增加将不一定导致群体数目的增加，因为数目不能超过生境的容纳量。可是，如果存在适合度的加性方差，则基因频率将改变，而且基因频率的改变可能造成与适合度相关的性状改变。因此适合度组分的改变是相关反应，一般，这比适合度本身的改变更有意义。因此我们必须考虑从原理上如何预测相关反应。

相关反应

如果一个性状与适合度相关联，那么对适合度的选择将在性状上产生明显的或相关的选择差。性状 Y 的相关选择差 S_Y' 是 Y 的加权平均值，权重是个体的适合度。因此：

$$S_Y' = \frac{\sum W(Y - \bar{Y})}{N}$$

通过类似于等式［20.1］的推导，这导致：

$$S_Y' = \mathrm{cov}_{P(YW)} \qquad [20.4]$$

即相关选择差等于性状和适合度的表型协方差。现在从等式［19.8c］即 $CR_Y/S_Y' = \mathrm{cov}_{(A)}/\mathrm{cov}_{(P)}$可以获得相关反应。根据这个式子和等式［20.4］，可以发现相关反应为：

$$CR_Y = \mathrm{cov}_{A(YW)} \qquad [20.5a]$$

即，性状 Y 和适合度的加性协方差。相应地，协方差可以用遗传相关和遗传力的形式来书写。这给出：

$$CR_Y = r_A h_Y h_W \sigma_Y \sigma_W \qquad [20.5b]$$

其中 σ 代表表型标准差。这样，要预测对自然选择的相关反应，必须知道性状和适合度之间的遗传相关以及性状和适合度各自的遗传力。Robertson（1966）涉及到奶牛中一个类似问题，就是从农场主采用的全面选择中预测相关反应，从而推导出等式［20.5a］。在自然选择下相关反应的详细论述参见 Crow 和 Nagylaki（1976）。

当对自然选择的相关反应作预测时，确保从一单个世代中获得适合度和性状值的估计值是重要的。例如，一般认为自然选择倾向于降低人的智力，因为来自较大家庭的孩子比来自较小家庭的孩子具有较低的智商；换句话说，对降低智商似乎存在一种相关选择差。可是在这种情形下，由于下列两个理由，关于选择差的假设是错误的：首先，因为具有 0 适合度的双亲没有孩子，因而不能出现在数据中；其次，因为它是以双亲适合度和孩子智商之间的相关为基础的。当比较双亲的智商与他们自己后来的孩子数作为适合度的一种度量时，发现相关的选择差是微小的正值（Waller，1971）。

选择的强度

自然选择对数量性状起作用的方式可从凭经验确定的适合度轮廓中推断出来。对于适合度的主要组分，像图 20-2 中的曲线（1），在其范围的大部分中选择是单向的。对性状（Y）起作用的选择强度可以根据相关选择差 S_Y' 和表型方差 $V_{P(Y)}$ 来表达。根据等式［20.4］，它表明这两个量的比率等于适合度对性状的线性回归，即：

$$b_{WY} = S_Y'/V_{P(Y)} \qquad [20.6]$$

这个回归系数被称为选择梯度（Lande 和 Arnold，1983）。如果同时考虑几个性状，那么每一个性状的选择梯度就是它的偏回归系数。对性状或物种之间的比较，S_Y' 要用性状选择前的表型标准差来除，做出按比例的调节（Endler，1986）。于是选择梯度是选择强度 i 的一种度量。选择强度告诉我们性状如何强烈地被选择，但是它不允许我们推断自然选择将引起的性状变化；等式［20.5b］表明预测反应将需要哪些其他参数。

像图 20-2 中曲线（2）那样，具有中间最优度的性状受稳定选择的控制，它是双向的。根据数据是否对方差或对适合度适用，所采用的稳定选择强度有两种度量。具有极端值（高或低）的个体是选择不利的。因此，与选择之前相比，选择之后由选择引起的性状表型方差更小。稳定选择强度可从由选择引起的同一世代的方差差异中来确定，差异被表达为选择前方差的比例。用 j 代表选择强度的度量，可通过下式来估计：

$$j = (V_P^* - V_P)/V_P \qquad [20.7]$$

其中，V_P 是选择前的表型方差，V_P^* 是选择后同一世代的方差（Lande 和 Arnold，1983；Endler，1986）。注意在稳定选择下 j 是一个负数。同样的度量可以应用于歧化选择，它增加方差；于是 j 是正值。

第二种度量以具有不同性状值的个体适合度为基础。作出的假设是适合度随性状与其最优值的离差平方而降低。假定 y 代表性状值偏离其最优值的、以标准差为单位的偏差，假设 $W(y)$ 代表具有性状偏差 y 的个体适合度。于是具有偏差 y 的个体相对适合度 $w(y)$ 是 $w(y) = W(y)/W(y=0)$。可以看出相对适合度对偏差平方的回归 b_{wy^2} 近似是线性的。这个回归的斜率就是稳定选择的梯度（Lande 和 Arnold，1983）。用 V_S 代表的选择强度的度量是适合度轮廓宽度的一种度量。当选择弱时，轮廓宽而扁平；当选择强烈时，它是窄而尖的。因此，V_S 的高值表示弱的选择，低值表示强烈选择。V_S 可从回归系数中获得：

$$V_S = -1/(2b_{wy^2}) \qquad [20.8]$$

为比较的目的，采用环境方差 V_E 的形式来表示。对于稳定选择，b 必然是负的，因此 V_S 是正的；对歧化选择，V_S 是负的。

以 V_S 的形式表示，稳定选择发生之后的表型方差是 $V_P^* = V_P - [V_P^2/(V_P + V_S)]$（Bulmer，1985，p.151）。因此 j 和 V_S 之间的关系是：

$$j = -V_P/(V_P + V_S) \qquad [20.9]$$

Turelli（1984），Endler（1986）和 Crow（1989）已经总结了自然群体中自然选择强度的估计值。强度是变化的，但经常在人工选择采用的同一范围内。稳定选择 V_S 的观察值很粗略地从 $100V_E$（弱的选择）变化到 $10V_E$（强烈选择），V_E 是性状的环境方差。强烈的稳定选择比强烈的歧化选择更普遍。真实的选择强度经常比它显现的要高，因为经常度量的仅是生活史中一个点上的选择。对单个性状自然选择的大小和方向估计还可能对自然选择的靶子和方式给出错误印象，因为选择是对适合度起作用的，因而是同时对很多表型上相关的性状起作用，但这些性状不全是可识别或可度量的。使用多元回归来估计多个性状上的定向选择和稳定选择的梯度（Lande 和 Arnold，1983）以及对整个生活周期中每一个阶段分剖这些梯度可部分地解决这些问题（Arnold 和 Wade，1984*a*，*b*）。尽管如此，很显然，很多性状具有落在图 20-2 的曲线（1）和（2）之间的适合度轮廓，并具有在偏离平均值某一距离处的最佳值。

平衡群体

现在我们可以看看自然选择的连续作用是如何令数量性状的遗传特性成形的。如果环境条件在一个很长时期内稳定不变，同时自然选择的方式和强度也是恒定不变的，那么群体将达到一种平衡，平衡时任一位点上的基因频率不再变化。因此所有数量性状的平均值保持不变。可能几乎没有实际群体严格处于平衡，因为环境不可能足够长时间地保持不变，但是对于下面的考虑，大多数群体可能足以接近平衡状态。然而，有一个限制条件：这里我们将作为新变异一种来源的突变忽略不计了。

适合度

根据费歇尔基本法则，很明显，当适合度的加性遗传方差耗尽时，适合度的选择反应

将停止。因此在平衡群体中适合度选择将没有进一步的反应，尽管在每一个世代都继续进行选择。也就是说，群体适合度已达到一种选择极限。在极限时保持的适合度遗传方差一定主要是非加性的，即由显性和上位互作引起的方差。这来自两个来源：来自对适合度为超显性的中等频率的基因，或来自因突变平衡选择而保持低频率的有害隐性基因。两种情形中，显性是定向的，因此预料适合度有严重的近交衰退。这已经重复地得到证实：果蝇（*Drosophila*）纯合染色体的适合度仅是杂合染色体的 10%（Sved，1971，1975；Mackay，1985*a*）。中等频率的基因比低频率的基因造成更多的变异，因此，可能是适合度的大部分方差来自超显性位点，并且大部分近交衰退来自稀有隐性基因（Crow，1952）。

因为对于适合度最大化而言，平衡群体中基因频率列阵最佳，所以，假若选择的是任一数量性状而不是适合度自身，并且假定存在选择反应，那么影响性状位点上的基因频率必定要变化，这是必然的。因此作为一种相关反应，适合度一定会降低，除非选择的性状完全由对适合度没有影响的基因所控制。这种预计已为经验充分地证实了：对数量性状的实验选择几乎总是导致适合度一种或多种主要组分的减少。仅提供一个例子：对果蝇（*Drosophila*）腹部刚毛数选择 5 个世代之后，估计了它的平均适合度作为竞争指数（Latter 和 Robertson，1962）。有两个向上选择的系和两个向下选择的系。相对于未选择的对照而言，在向上选择的系中平均适合度是 79%，在向下选择的系中为 65%。

如果人工选择得以进行，然后在很多变异由于固定而丢失之前中断，那么自然选择一定倾向于使基因频率恢复到它们的平衡值，并且预计人工选择的性状平均值回复到它的初始值。自然选择抵制基因频率变化的这种倾向被称为遗传的自动调节（Lerner，1954）。当加权选择差低于未加权的选择差时，实验选择过程中能经常看到它的效应（见例 11.5）。

主要组分

对经常称为生活史性状的适合度组分的持续自然选择与对适合度本身的选择具有类似结果：加性遗传方差减少，导致低遗传力，并且继续存在的非加性方差归因于定向显性，导致近交衰退。Kearsey 和 Kojima（1967）比较了果蝇（*Drosophila*）的 12 个性状。适合度主要组分的所有度量都表现出上位互作和强烈的定向显性，而所有的其他度量即身体大小和刚毛数的度量都表现出小的或没有显性和互作。在用果蝇做的另一个实验中（Mackay，1985*a*），获得了存活力和生殖力的适合度轮廓图，它们与图 20-2 中的曲线（1）极其相似。这些性状也表现出大量的近交衰退。Roff 和 Mousseau（1987）以及 Mousseau 和 Roff（1987）已经总结了普遍证实适合度组分低遗传力的广泛资料。

如果考虑加性方差本身而不是遗传力的数量，那么出现的适合度组分图象稍有点不同。一项实验调查（Houle，1992）表明如果方差按比例调节到平均值上，即表达为变异系数，那么适合度组分比其他性状具有更多的加性方差。低遗传力似乎是由大量的非加性和环境方差而不是由少量的加性方差引起的［见 Houle（1992）的讨论］。

适合度组分和适合度本身之间的差异是最高值的单独组分没有最高的适合度。适合度可能作为一种指数来考虑，通过它自然选择可同时选择所有的主要组分。于是，我们应当预料作为适合度主要组分的性状之间的加性遗传相关是负的，其理由已给出在前一章的末尾。作为一个例子，考察图 20-1 中的出生数目。在范围的低端，出生的每一个额外的仔畜几乎产生一个达到成年的额外后裔，因此总适合度增加了。但随着出生数目的增加，由

于哺育能力的限制它们的“质量”逐渐降低，直到超过一确定数目，降低的质量超过额外的数目，几乎没有几个后裔存活到成年。这类负的加性遗传相关，有时称为“颉颃的基因多效性”，对随机交配的果蝇群体的适合度组分已经发现了（Rose 和 Charlesworth，1981）。因此，适合度组分的加性遗传方差可通过与其他适合度组分的负遗传相关来保持。可是，在近交系中，预计适合度组分之间的遗传相关是正的，因为经近交而纯合的隐性有害基因很可能对一种以上的适合度组分具有害影响，因此对遗传相关有正的贡献。实际上在上面引用的实验中，Mackay（1985a）发现了高的正相关。

如果一个平衡群体适应的环境变化了，那么基因频率的阵列不再是最好的。变化了的环境条件更改了适合度组分的相对权重，以致于现在适合度自身有了一些加性方差，因而对自然选择会有反应。人工选择的应用可被认为是以这种方式改变了环境条件，更改了联合的自然和人工适合度中的组分权重。由于医学和节育、心理和行为因素已大量代替了生理因素作为适合度决定因素的结果，使得人类群体适合度的组分在最近发生了剧烈的变化，特别是在富裕的国家。

具有中间值最优的性状

在一种纯粹描述意义上我们指出中间值最优适合度轮廓是稳定选择的结果，这只是意味着中间表型具有最高的适合度。可是，一个中间值最优的观察并不告诉我们自然选择如何对性状或对影响性状的基因起作用。为什么中间型可能更适应，有若干不同的理由（Robertson，1967），这会产生关于平衡群体中这些性状的遗传组成的不同预测。

“真实”稳定选择 稳定选择的典型模型是性状值为适合度的一个直接原因。我们将这种情况称为“真实”稳定选择，暗示着性状值对适合度的因果关系。一个易懂的对适合度有直接原因效应的性状实例是对哺乳动物皮毛隔热的任何度量。不活动期间贮存热量和活动期间散热，这种互相矛盾的需要以中等程度的皮毛密度来保持均衡。中间个体是有利的，因为性状值最佳。选择是功能性的或真正的稳定选择。（据说这个例子是猜想的：如果进行实验，可能发现在任何实际群体中对变异的整个范围来说，与性状有关的适合度的差异是相当小的。）Robertson（1967）表明真实的稳定选择不象是普遍的，因为模型基本上将生物体细分成与可想象得到的性状一样多的分区，每一个分区通过它的平均表型值作为原因来影响适合度。

真实的稳定选择有两个主要效应。首先，不管基因是否影响性状的平均值，它有利于变异性最少的基因型（Curnow，1964）。（第 19 章中基因型 × 环境互作的考察表明稳定选择降低环境敏感性。）换句话说，它增加发育的限向作用（canalization）（Waddington，1957）。其次，它倾向于减少性状的遗传方差，它以两种方式做到这一点（Bulmer，1971，1976）。更直接的开始最大的效应是通过配子期不平衡而产生的。选择造成单个基因型中不同位点上负相关的等位效应。因此等式［8.10］中协方差一项是负的，遗传方差是下降的。就连锁的基因来说，选择倾向于建立“平衡”的组合，其中连锁的基因主要处于相斥连锁，这样染色体作为一个整体的效应被最小化了（见 Mather，1941；Lewontin，1964）。可是，选择必须很强烈，或连锁很紧密以保持组合中的基因明显地不同于随机排列的基因（见 Wright，1969，p.92）。改变影响性状平均值的位点上的基因频率是减少遗传方差的第二种方法。只要基因型在变异性方面没有很大的不同，并且只要位点除了通过性状不以任

何其他方式影响适合度，那么稳定选择就倾向于将基因频率改变到固定，因此减少遗传方差（Robertson，1956）。可是，基因频率改变很慢，除非选择强烈或者基因有大的效应。因此真实稳定选择的结果是既减少环境方差又减少遗传方差。在实验中这两种效应已观察到；例如，果蝇（*Drosophila*）腹侧刚毛数（Gibson 和 Bradley，1974），和赤拟谷盗（*Tribolium*）的蛹重（Kaufman，Enfield 和 Comstock，1977）。

"表观"稳定选择　数量性状的中间值最优产生于影响性状适合度组分的基因多效性效应。例如，生殖力的中间值最优可能由后裔的生殖力和存活力之间的负相关产生。鸟类的窝卵数是一个众所周知的中间值最优的例子，可以用这种方式解释（Lack，1966）。窝卵数（产卵数）应该是适合度的一种主要组分，但是可被养育的后代数目受到可利用的食物供给限制。因为适合度两种主要组分幼仔数目和质量之间的平衡，中间的窝卵数导致养育的最大数目。小鼠身体大小的中间值最优可能类似地由适合度不同组分的相反相关引起。身体大小与出生数目呈正相关，因为大的鼠具有较大的窝产仔数（Falconer，1965*b*），但与存活力可能呈负相关，因为大的鼠安静，对干扰不起反应，而小鼠活泼，对干扰起有力的反应（MacArthur，1949；Falconer，1953）。在自然条件下，人们必须假定较大的鼠与小鼠相比将不擅长于逃避食肉动物。同样地，果蝇（*Drosophila*）身体大小的中间值最优是由生殖力、发育时间和幼虫存活力的相反遗传相关产生的（Roff，1981）。所有这些例子中，稳定选择是假设的，因为具有中间值最优的性状不是选择的直接目标。

黑腹果蝇（*Drosophila melanogaster*）腹侧刚毛数为与性状自身无关的来自基因多效性效应的表现稳定选择提供一个清晰的实例（Kearsey 和 Barnes，1970）。在成年果蝇胸的侧面上，腹侧刚毛是小刚毛。研究的群体是从反向选择的两个品系的杂交中衍生的；雌性中平均有 18.5 根刚毛，变幅约为 10～45。稳定选择的证据给在图 20-3（*a*）中，该图显示了在不拥挤条件下（实线）和拥挤条件下（虚线）发育的果蝇刚毛数的分布。在拥挤条件下，对幼虫存活的自然选择要强烈得多，可明显看出它有消除具极端刚毛数果蝇的效应。从有不同刚毛数果蝇的频率下降中可以构建适合度轮廓图。可是，这个轮廓不是适合度本身的，而是一个主要组分即幼虫存活的。图 20-3（*b*）显示了这种轮廓图，在平均刚毛数处有适合度的尖状峰。关于这种适合度轮廓的重要点是在果蝇长出任何刚毛之前，引起中间值最优的选择发生在幼虫阶段。因此，具有中间刚毛数的果蝇其优良适合度决不是由性状本身引起的，而必定是对刚毛有多效性效应影响幼虫存活的基因所引起。稳定选择是假设的，就像以前例子一样，但是在这种情形下，我们不知道选择的真实标准是什么，也不明白为什么具有中间刚毛数的成年个体在幼虫阶段是最适合的。

不知道选择的是适合度的什么组分以及观察的性状与这些组分的相关如何，就不能预测表观稳定选择对性状遗传方差的影响。一种可能性如下（Robertson，1956）。影响适合度主要组分的基因可能对讨论中的性状具有基因多效性效应，对性状的这些效应可能或多或少是加性的。于是，具有性状中间值的个体与极端个体相比，在较多位点上一定是杂合的。如果位点对适合度具有超显性效应，那么自然选择有利于杂合子，结果显然有利于性状的中间个体。以这种方式起作用的表观稳定选择的结果将是性状遗传变异的保持。可是，在上面描述的腹侧刚毛数的情形中，表观稳定选择不是杂合子优势的原因，因为对纯合品系也观察到相同的适合度轮廓（Linney 等，1971）。Kimura（1983）提出如果在稳定选择下控制性状的位点数很大，那么在每一个位点上起作用的选择一定很小。于是在这些位

点上的等位变种将很近于选择中性的，即使性状表型被选择，它们的基因频率也是由非频发突变和漂变之间的平衡来确定。这能够解释表型水平上自然选择的优势与分子水平上中性或几乎中性等位基因的优势之间的矛盾。

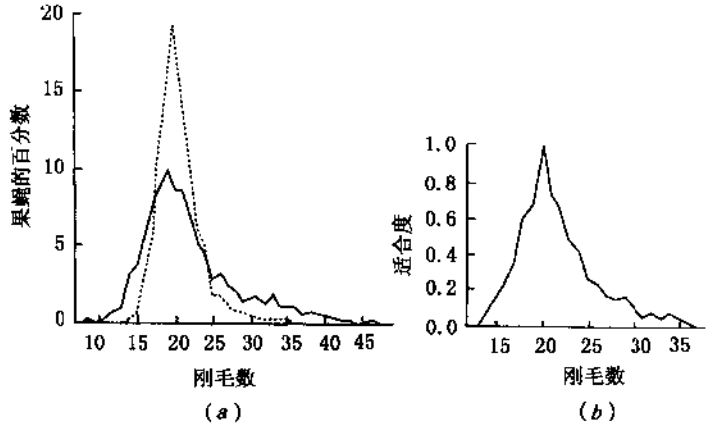


图 20-3 (*a*)果蝇腹侧刚毛数的分布:连续直线,在弱自然选择下;虚线,在强自然选择下。(*b*) 从 (*a*) 图频率分布中获得的适合度轮廓图。适合度的尺度相对于最适合的刚毛等级。度量的适合度仅是从卵存活到成虫的组分，不是适合度的全部（根据 Kearsey 和 Barnes，1970 改编）

具有中间值最小适合度的性状

歧化选择是稳定选择的对立面：选择不利于中间个体。虽然这不经常发生，但在自然群体中已观察到（Endler，1986）。最可能的原因是存在某个方面不同的可选择的栖息地，这使得一个表型在一种栖息地极端有利，另一个在另一种栖息地极端有利。在加拉帕戈斯群岛上的地面雀类提供了一个与不同栖息地利用有关的歧化选择的实例。具有不同喙形状和大小的鸟在捕获和剥食不同大小的种子方面更有效。具有小而狭窄喙的鸟吃小而软的种子更有效，面具有深喙的鸟利用大而硬的种子更有效；这导致中间大小喙的鸟相对于极端的鸟有下降的适合度（Schluter，Price 和 Grant，1985）。预期歧化选择同时增加遗传和环境方差，这些效应在实验中观察到了。例如，当应用到果蝇幼虫发育时间上时（Prout，1962），环境方差增加了；当应用到赤拟谷盗的蛹重上时（Halliburton 和 Gall，1981），遗传和环境方差都增加了。

中性性状

很多数量性状的适合度具有像图 20-2 中曲线（3）那样的轮廓，能否说这些性状就真正是中性的尚不清楚。人类的指脊数可能是这样一个例子。注意，说一个性状的适合度是中性的，并不意味着性状没有功能，面是指性状的精确值不是适合度的一个决定因素。一些性状确实具有很广泛的最优值，超出大多数观察范围的性状值变异与适合度的差异没有关联。确定一个性状的变异是否是中性的一种间接方法是进行“干扰”实验。一个群体经过几个世代对性状进行定向选择；然后，当平均值改变到偏离其初始值达某种程度时，终止选择，随机繁育群体，仅接受自然选择。如果平均值不回复到它的初始值，或者只是很慢地回复到它的初始值，那么可以得出结论，性状是中性的，或近乎是中性的。严格地

说，也应当应用反向选择，以证明平均值可被恢复。果蝇（*Drosophila*）腹部和腹侧刚毛数的干扰实验表明这些性状可能是很近于中性的（Latter 和 Robertson，1962）。这个结论为经验上确定的染色体纯合子腹侧刚毛数的总适合度轮廓所支持（Spiers，1974；Mackay，1985*a*）。另一方面，上面描述的对腹侧刚毛数的表观稳定选择表明影响刚毛的基因确实也影响适合度，尽管影响不经过刚毛数目。另外，自然群体中刚毛数在时间上的和空间上的稳定性表明涉及的基因就适合度而言不是中性的。刚毛是感觉器官，这些基因的主要功能可能通过外周神经系统的发育来影响适合度。只有当我们知道基因的主要功能是什么的时候，才可获得大多数数量性状（即不是适合度主要组分的性状）是如何与适合度相关联的完整的理解。

突变引起的变异来源

自然群体中的自发突变产生新的遗传方差，在前面对自然选择如何影响数量性状的解释中这点被忽略了。选择降低方差，突变增加方差，仅歧化选择除外。在下面的部分，我们将不得不考虑选择和突变之间的平衡能否解释在自然群体中发现的遗传方差的数量。首先，我们必须描述关于突变产生新变异的速度我们知道些什么。

突变方差

在一个世代中由突变产生的任一性状的方差叫做*突变方差*，用符号 V_M 表示。为了有利于在性状之间和在物种之间进行比较，V_M 被表示为性状环境方差 V_E 的一个比例。有两种方法可用来估计 V_M。我们在第 12 章和第 15 章看到，新突变对长期选择反应以及近交系亚系之间的趋异有贡献。因此，通过使一个群体纯合，并将它分成许多在全部时间里都积累突变的重复亚系，可以估计 V_M。如果对亚系有关性状进行了选择，那么反应的速度提供 V_M 加性组分的一个估计值。如果没有对亚系进行选择，V_M 可从它们之间性状平均值的趋异中来估计。两种方法仅估计出由中性突变引起的新方差，因为有害突变体被自然选择淘汰了。采用两种方法，在大的重复实验中已获得了果蝇（*Drosophila*）腹部和腹侧刚毛数 V_M 的几个估计值（这些实验的一个综述以及早期研究，见 Keightley，Mackay 和 Caballero，1993）。果蝇刚毛性状 V_M 的估计值一致地为 10^{-3} 乘以环境方差 V_E，即 $(V_M/V_E)=10^{-3}$。对果蝇生活史性状（Houle 等，1994）和酶活性（Clark，Wang 和 Hulleberg，1995）以及小鼠身体大小（Hill，Caballero 和 Keightley，1994）都报道了类似水平的突变方差。其他物种中其他性状的平均突变方差也是 $10^{-3}V_E$，具有从 10^{-4} 到 $10^{-2}V_E$ 的范围（Lynch，1988）。

如果性状的遗传力已知，则我们可以回答下列问题：对突变来说，将花费多长时间才能产生在自然群体中发现的遗传方差？例如，果蝇（*Drosophila*）刚毛性状具有大约 0.5 的遗传力，有小的非加性方差。因此 $V_G=V_E$，每世代 $V_M=10^{-3}V_G$。因此，如果基础群体没有遗传方差，对突变来说，将花费1 000个世代才能产生自然群体的遗传方差。

关于组成 V_M 的单个参数的大小，我们了解的还不多。如果所有位点上的突变等位基因对性状具有相同的效应并具有相同的突变率，那么 $V_M=2nua^2$，其中 n 是位点数目，该位点的突变影响性状；u 是每一个位点的突变率；a^2 是突变等位基因效应的方差（假定

在每一个位点上突变的平均效应是 0，因此，增加和降低性状值的突变经常是均等的）。与这里有关的 n 是位点的总数，在这些位点上，影响性状的突变能够发生，n 一定比对选择反应有贡献的位点数（第 12 章）大，或一定比对自然群体中分离变异有贡献的位点数（第 21 章）大。影响数量性状的位点上的突变率不可能低于具有质量突变体表型的位点上的突变率（10^{-5}到 10^{-8}），无疑地要高得多。具有数量效应的突变可能位于影响表达的时间控制或蛋白质活性的基因调节区域，而质量效应的突变很可能位于外显子（exons），并影响基因产物的数量和作用；因此对数量效应的突变来说，其突变靶子的大小可能较大。唯一的直接证据来自度量每个性状突变率（nu）的实验。Mukai 等（1972）估计了影响果蝇存活力，即从卵存活到成虫的自发突变的起源速率。新突变在源于一单个祖先的复制的第 2 号染色体上积累了 40 个世代，通过一个平衡器染色体防止了突变遭受自然选择和重组。这些染色体对幼虫存活的纯合效应以 10 个世代的间隔来测验，从中推断出第 2 号染色体上突变的起源速率。有两类突变基因：一类突变基因在纯合子中是致死的或严重有害的，一类是适度有害的，具有的存活力大约是正常个体的 0.6 以上。40 个世代之后，24% 的染色体携带一个或多个致死因子，起源速率是每世代每条染色体 0.6%。有害突变大约是经常发生的 10～20 倍，具有每世代每条染色体大约 10% 的起源速率。通过 5/2 按比例放大来代表整个基因组，这意味着一个世代的突变之后，大约 25% 的配子进行一次新的适度有害的突变；即 $nu = 0.25$。类似地，小鼠骨骼变种和玉米形态性状的每性状突变率是每世代每一个配子超过 0.01〔Turelli（1984）综述〕。nu 的这些高值确实表明有大量的突变位点影响数量性状，具有数量效应的等位基因有高的突变率，或者两者都存在。注意，由于每个基因组位点数目对 n 所设定的上限，每个性状大的 n 暗示着新突变对多个性状具有广泛的基因多效性效应。

影响数量性状的新突变的纯合（a）、杂合（d）以及基因多效性效应的信息主要来自研究影响果蝇（*Drosophila*）存活力和刚毛数的自发和诱导突变。正如上面提到的，影响存活力的突变其纯合效应从致死变化到准正常。平均而言，杂合存活力的效应对两类突变是相似的，近似为 0.02，但是，在显性度和效应的严重性之间存在一种相反的关系，致死突变几乎完全是隐性的，适度有害突变大致是加性的（Crow 和 Simmons，1983）。在 Caballero 和 Keightley（1994）综述中以及 Mackay、Lyman 和 Jackson（1992）用 P 因子进行的插入突变等其他实验中，发现了对刚毛性状的类似影响〔P 因子是黑腹果蝇（*Drosophila melanogaster*）的转座（位）因子，它们引起易位后它们插入的基因发生突变〕。从这些实验中获得的结论如下：(i) 大多数突变对刚毛数有很小的效应，但具有巨大效应的少量突变也出现了。突变效应的分布呈瘦峰状，来自新突变的方差增加其大部分是由巨大效应的那些突变引起的。(ii) 影响刚毛数的新突变的显性度存在着变异，巨大效应的突变具有倾向于隐性的一种趋势。(iii) 对刚毛数具有巨大纯合效应的突变还降低纯合的存活力。对 P 因子插入突变，在刚毛数和存活力上效应的绝对值相关对腹部和腹侧刚毛数都大约为 0.4。

遗传变异的保持

自然选择倾向于降低数量性状的遗传变异，有限群体中的随机漂变也是这样，然而，在实践中我们观察到自然群体中所有性状都仍有遗传变异。自发突变能够抵制这些倾向并

足以保持观察到的遗传变异水平吗？或者我们必须求助于其他的平衡力量来对诸如适合度的基因型×环境互作作出解释吗？这个问题在理论上已有相当广泛地阐述（见 Bulmer，1989 以及 Barton 和 Turelli，1989 的综述）。准确的答案取决于（i）适合度轮廓的形状和选择强度，（ii）有效群体大小，（iii）突变方差，这又取决于位点数目、突变效应的分布、每一个位点的突变率。我们已经看到，很多参数的值是未知的，或仅仅很粗略地知道。数量变异保持的理论模型必须与目前已知的这些变量的观察值相一致。用简单的模型难于满足这种要求，但提出的问题指明了将来经验和理论研究的方向。

中性突变和随机漂变之间的平衡

在第 15 章我们看到，当新突变方差被随机漂变造成的损失所抵消时，那么在有效大小为 N_e 的群体中现存的遗传方差是 $V_G = 2N_eV_M$。表型方差是 $2N_eV_M + V_E$，因此遗传力是：

$$h^2 = \frac{2N_eV_M}{2N_eV_M + V_E}$$

这是广义遗传力，因为 V_M 包括所有的遗传方差，不仅仅是加性的。如果我们代入 $V_M = 10^{-3}V_E$，则遗传力变为：

$$h^2 = \frac{0.002N_e}{1 + 0.002N_e}$$

对不同群体大小，这个算式给出下列的值：

$N_e = 100$	1 000	10 000
$h^2 = 0.17$	0.67	0.95

（见 Lynch 和 Hill，1986）。因此，除很小的群体之外，在没有选择的情况下，突变能够保持大量的遗传变异。一些性状诸如人类指脊数（见第 10 章）的很高遗传力表明在对影响它们的基因没有选择时，它们受控于中性突变。另一方面，大多数性状的遗传力比通过中性模型预测的值要低得多，我们必须得出结论：对于大多数性状，在降低遗传方差方面选择也起部分作用。

突变—选择平衡

在第 2 章和第 4 章中我们讨论了保持多态性的机制。现在我们必须将这种讨论扩展到其变异是由多态位点引起的数量性状。为此，我们必须考虑突变方差和选择强度的观察值，弄清楚这两种力量的平衡是否能解释观察到的遗传方差水平。

适合度性状　对于作为适合度主要组分的性状，在自然群体中观察到的遗传变异水平与从突变和自然选择之间平衡中预测的变异是一致的。影响这些性状的大多数突变等位基因是有害的，并且选择是单向的——不利于突变体。可以证明，新的有害突变的输入和不利于它们的定向选择之间平衡时预期的遗传方差近似为 $V_G = V_M/s$，其中 s 是不利于杂合突变体的选择系数（Barton，1990）。关于果蝇（*Drosophila*）卵到成虫存活力的实验已经估计出 $V_G = 1 \times 10^{-2}$，$V_M = 1.3 \times 10^{-4}$，以及 $s = 0.02$〔Mukai 等，1972；Mukai（1979）的综述〕。因此预期的平衡遗传方差近似为 0.65×10^{-2}，这与观察值有着良好的一致性。

"真实"稳定选择　对于处于真实稳定选择之下的性状，当突变不是绝对有害时，则情况更复杂。有两种主要模型，关于它们仅能给出要点。第一种模型（Lande，1975）假

设在 n 个位点的每一个位点上存在无数个可能的等位基因，并且对于每一个位点，在平衡群体中分离等位基因的效应和突变等位基因的效应都是正态分布的。由这种模型预测的平衡方差是：

$$V_G = \sqrt{(2nV_MV_S)} \qquad [20.10]$$

其中、V_M 是突变方差，V_S 是上面解释的稳定选择强度。只有相对于每一个位点上现有的等位基因变异的突变等位基因效应方差很小，关于平衡时正态分布的等位基因效应的假设才是合理的近似。这种要求意味着只有当突变等位基因的效应都小，因而稳定选择弱时以及当每个位点的突变率很高时，模型才是有效的（Turelli，1984）。为了弄清楚这种模型是否能解释观察到的遗传方差水平，我们考察经受稳定选择的性状的高达 0.5 的遗传力并回答是否与 $h^2 = 0.5$ 相一致。我们取观察值 $V_M = 10^{-3}V_E$，对弱选择 $V_S = 100V_E$，看看由模型包含的位点数是否是合理的。于是 $V_G = \sqrt{(2n \times 10^{-3} \times 100)}V_E$。当 $h^2 = 0.5$ 时，$V_G = V_E$，这给出 $n = 5$。这意味着只有 5 个突变的同等位点对于产生观察到的遗传变异数量是必要的，这似乎是合理的。可是，如果只有几个位点，那么每个位点的突变率必须很高才会产生观察到的突变方差，或等位基因效应必须巨大，这将与模型不一致，因为等位基因效应的分布不再是正态的。因此模型似乎不能解释观察到的遗传方差的数量。

第二种模型（Turelli，1984）也假设新突变的效应是正态分布的，但是突变效应与影响性状的每一个位点上的现有等位基因变异相比要大；每个位点的突变率比 10^{-5}要低得多；选择强度 V_S 在观察到的 $100V_E$ 到 $10V_E$ 的范围内。于是平衡遗传方差是：

$$V_G = 4nuV_S \qquad [20.11]$$

在这些条件下，平衡时等位基因效应的分布呈高度瘦峰状，大部分方差由具巨大效应的稀有等位基因来贡献。为了评价这种模型，我们假设 $V_S = 10V_E$ 的强烈选择，我们找出当 $h^2 = 0.5$ 时，乘积 nu 的预测值是多少。这样 $V_G = 40nuV_E$，$nu = 2.5 \times 10^{-2}$。为了保持可估计的遗传力，要么 n 要么 u 必须很大。如果我们取 $n = 100$ 作为合理的位点数，那么每个位点的突变率必须为 2.5×10^{-4}，这比认为是合理的突变率至少要高出一个数量级的大小。如果我们取 $u = 10^{-5}$，10 倍的可接受的主要突变体突变率，那么 n 必须为2 500,这显然是一个不合理的巨大的位点数目。如果 n 很大，那么一定存在着普遍的基因多效性。如果一个突变等位基因对一个性状具有巨大的效应，而对很多其他性状具有较小的基因多效性效应，那么位点上保持的方差将由对等位基因具有巨大效应的性状所做的选择控制，因此位点对等位基因效应很小的其他性状的遗传变异的贡献将比预测的要小得多（Turelli，1985）。

正如上面提到的，真实稳定选择的概念要求选择设想生物体被细分成与可想象的性状一样多的分区，每一个分区在原因上影响适合度。简单的遗传负荷论据表明，具有单个性状观察到的数值大小的真实稳定选择不能独立地很对多性状产生作用（Robertson，1967；Barton，1990）。真实稳定选择还意味着新突变的一个显著的部分将是有益的，新突变对适合度的影响取决于它发生的遗传背景。关于方差普遍可接受的观点是，大多数突变是绝对有害的或充其量是近乎中性的（Kondrashov 和 Turelli，1992）。通过这种直接机制来保持数量性状观察到的遗传变异的水平似乎是不大可能的。

“表观”稳定选择 如果突变对适合度有有害的基因多效性效应，那么稳定选择的表

观会出现，因为具有极端表型的个体是较不适应的，如果它们有更多的有害突变（Barton，1990）或者如果对性状有巨大效应的突变对适合度也有巨大的效应（Keightley 和 Hill，1990）。数量遗传变异可通过突变与不利于新突变适合度的有害基因多效性效应的定向选择之间的平衡来保持，这一思想是吸引人的，因为它考虑了观察到的突变特性。可是，用简单的基因多效性模型来同时解释观察到的分离变异的水平和自然选择的强度仍然是困难的。例如，给定具有杂合适合度为 $1-s$ 的同等有害突变，以及相关于突变的强烈选择，则平衡遗传方差是 $V_G = V_M/s$，表观稳定选择强度近似为 $s = V_M/V_G$（Barton，1990）。因此，如果 $V_M/V_E = s \approx 10^{-2}$，则遗传力为 0.5 和强烈的稳定选择（$s = 0.02$）是可能的。$V_M/V_E$ 的估计值要比不利于新突变杂合效应的平均选择的估计值低一个数量级。因此，通过假定影响数量性状的突变是近中性的（$s = 10^{-3}$），借助这种模型，人们能够解释高水平的分离变异，但这又不能解释观察到的稳定选择的强度。

如果影响性状的突变是有害的，则获得的突变方差的估计值是向下偏倚的；这种偏差是否多达一个数量级的大小取决于突变对性状和对适合度效应的联合分布。对于保持具有中间值最优的性状的遗传方差，突变—选择平衡模型有效性的进一步评估将取决于这种分布以及引起数量变异位点上的突变率的更准确的经验测定。如果新突变的频率和效应的实验值不是所要求的大小，那么必须探索可供选择的解释。例如，可能的说法是，影响数量性状的突变效应在不同的环境中有可能不同。如果这样，时间和空间的环境变异能够保持遗传变异，因为平均而言杂合子将比纯合子更具有适应性（Barton 和 Turelli，1989）。

习　题

20.1　习题 11.2 与达尔文雀科的一个物种进化变化有关，该物种遵循关于喙深的选择生存。根据这个假设，即“选择生存的原因是喙深本身，不是与它相关的某一个其他性状”，从 $R = h^2S$ [11.2] 可预测这种自然选择的反应。证明如果：

$$\frac{r_A}{r_P} = \frac{h_Y}{h_W},$$

预测将是有效的。其中 Y 是喙深，W 是适合度，相关是喙深和适合度之间的。

[解答序号：2]

20.2　从下面有关明尼苏达白种人一个样本的数据中可以获得自然选择对智商得分影响的大致概念。数据涉及到个体的智商得分以及这些个体完整家族的大小，即他们的孩子数目。必须假定家族大小是适合度的一种充分的度量。平均值、标准差和遗传力是

	平均值	标准差	遗传力
智商得分	103	15.4	0.6
家族大小	3.4	2.3	0.1～0.2

家族大小的遗传力没有可靠地估计，但可能在标明的范围内。两个性状之间的相关是 +0.11。必须假定遗传相关与表型相关没有差异。以这些数据和假设为基础，预测每个世代智商得分的变化是多少？这个群体中有关智商得分表观的，即相关的选择差是多少？

数据来自 Waller, J.H.（1971）*Social Biol.*，**18**，122～136。

[解答序号：65]

20.3 为下面人的出生重数据作一个像图 20-2 那样的“适合度轮廓”图。分析了1974年在意大利出生的所有婴儿的纪录。这里的数据涉及正常妊娠 9 个月之后出生的男婴，共有413 572个男婴。对出生重进行分组，并给出了它们的中点。平均出生重是3.46kg，标准差是 0.51kg。对每一个出生重组，给出了头 4 周每千人的死亡率，包括死产。为作适合度轮廓图起见，必须假定存活到 4 周龄等价于适合度。

出生重（kg）	频率（%）	每千人死亡数
1.3	0.13	612
1.8	0.35	333
2.3	2.13	94
2.8	15.95	27
3.3	40.32	15
3.8	30.73	11
4.3	8.54	12
4.8	1.56	25
5.55	0.30	69

数据来自 Terrenato，L. 等．（1981）*Ann. Hum. Genet.*，**45**，55～63，和 Ulizzi，L. 等．（1981）*Ann. Hum. Genet.*，**45**，207～212。

[解答序号：139]

第 21 章　数量性状位点

在前面各章中描述数量变异的遗传基础不得已地只考虑了引起变异的所有基因的总效应。一个完整的描述需要逐个地考虑基因的特性即它们的基因频率以及它们对有益性状影响的大小。不能采用经典的孟德尔遗传学方法来逐个研究基因，因为它们的效应丢失在所有其他背景变异的统计“模糊”中。在对基因的单个特性缺乏了解的情况下，理论研究工作不得不作出一些不切实际的假设，例如，所有位点上的基因频率到处是相同的，基因的效应和显性关系也大约是相同的，并且在某些范围内存在无数多个影响性状的基因。可是，最近已有方法变得可用于研究单个基因；这些基因被称为数量性状位点或 QTL（Quantitative Trait Loci）。这一章将说明识别 QTL 以及估计它们对数量性状影响的方法。而且，我们将看到，作为一个 QTL 被识别的是影响性状的染色体的一个片段，不必是一单个位点。

识别单个基因能带来一些有用的应用。首先，它能够改进选择性育种的功效，特别是对低遗传力或只能在一个性别中度量的性状（见 Soller 和 Beckmann，1988；Lande 和 Thompson，1990）。其次，转基因技术可以应用于数量性状。第三，在医学上，引起共同多因子疾病例如心脏病或糖尿病倾向的等位基因识别可以导致预防方法的改进。第四，当基因的数目和特性已知时，数量遗传学理论将变得更切合实际，进而将提高我们对进化的理解。

主效基因

凭借对一些等位基因的偶然发现，这些等位基因对性状的效应大到足以靠它们的单个分离可以识别的程度，而意外地识别出许多影响数量性状的位点。这包括，如在果蝇（*Drosophila*）、小鼠和玉米中经自发或诱发突变而产生的经典遗传学的主要的形态突变。已在选择品系中发现某些效应巨大的基因是分离的。这些实例有：粗糙的（scabrous）、截毛（bobbed）和盾片（scute）刚毛位点的等位基因，它们是在选择刚毛数的果蝇（*Drosophila*）品系中发现的（见 Mackay，1989 的综述）；侏儒位点上的等位基因（例 7.1），即肥胖和矮化等位基因，它们是在对小鼠体重做选择的品系中发现的（见 Roberts 和 Smith，1982 的综述）以及一个 *hg* 基因，它可以引起经增重选择的小鼠品系在断奶后快速增重（见 Bradford 和 Famula，1984）。表 21-1 列出家畜中对商用性状有主效应的等位基因实例。许多情况下在中选品系中可能因突变重新产生效应巨大的等位基因（第 12 章）。这样的主效等位基因，在未选择群体中频率很低，并且对数量性状的分离没有贡献。这几乎一定可归因子它们对适合度的有害的多效性副作用。纯合截毛（bobbed）突变的雌性果蝇（*Drosophila*）很少能存活，并且比野生型有更长的发育时间（Lindsley 和 Zimm，1992）；当小鼠的矮化、肥胖和高速生长基因纯合时，小鼠或不育或生殖力降低（Roberts 和 Smith，1982；Bradford 和 Famula，1984）；猪的氟烷（halothane）基因和牛的双肌（double muscling）

基因降低生存力和生殖力（Hanset，1982；Webb 等，1982）。

表 21-1 影响数量性状的主效基因

物种	性状	基因	参考文献
禽	身体大小	矮小	Merat 和 Ricard（1974）
猪	瘦肉，猪应激综合症	氟烷敏感	Smith 和 Bampton（1977） Webb 等（1982）
牛	瘦肉，肌肉肥大	双肌	Rollins 等（1972） Hanset 和 Michaux（1985*a*，*b*）
绵羊	多产性	Booroola F	Piper 和 Bindon（1982） Piper 等（1985）

如果存在对数量性状有巨大效应的等位基因，那么确定是否自然群体中有另外的效应巨大但对适合度效应无害的等位基因在分离，或者对选择反应有贡献，是有相当意义的。注意的是，这个意义上的“巨大”并不一定意味着是按描述为典型孟德尔突变所需的超过群体平均值 3 个标准差的效应，在这个范围内，0.5～1.0 σ_p 的效应便是“巨大”。例如，在频率 $p=q=0.5$，有一个表型标准差（$2a=1.0\ \sigma_p$）的标准化效应时，归因于一个严格的加性等位基因的加性遗传方差，是表型总方差的 12.5%（等式［8.5］）。但是，因为其他位点上变异的混杂效应，使得分离群体中有这种效应的一个等位基因将不很明显。

检测方法

有几种发现影响数量性状主效基因的方法；这里仅概括地介绍如下。

（1）多峰分布 如果相对于背景遗传的和环境的变异一个基因具有足够大的效应（超过大约 3 σ_p），那么它将在分离群体中产生一种多项分布。寻找多峰（multimodality）的最好方法是在两个异化品系之间杂交衍生的世代中。如果等位基因是部分显性的，则 F_2 中的分布将是三峰的，与更隐性的品系杂交时，其分布是双峰的；F_1 将显示这是哪一个品系。

（2）有选择的回交 这是一种通过降低背景遗传变异使回交杂种中双峰分布更清晰的一种方法。它是采用的最早的方法之一，是由 Wright（1952）提出的。通过选择（或近交）产生的两个异化品系杂交，并与更隐性的品系进行重复回交。在每一个世代中，对更显性表型的个体进行选择；即，如果是与低品系回交，则对高的表型进行选择。在选择方向上选择保持有巨大效应的任何等位基因分离。它的效应通过来自低品系的等位基因的杂合子对纯合子的双峰分布来检测。同时，在回交的每一世代中来自高品系的其他有小效应的基因频率都减半（表 5-2），因此背景遗传变异降低了，使得双峰分布更清晰。这种方法就是在 Booroola 美利奴绵羊培育过程中采用的方法，虽然是无意的，在此过程中有巨大多产性效应的等位基因 F（表 21-1）分离出来了（Piper 和 Bindon，1988）。

（3）非正态分布 一个基因，它的效应不足以引起多峰分布，但仍然可以引起可检测的偏离正态性的离差。如果基因的频率是中等的，则分布将是扁形峰状的（比正态更扁平）；如果它是极端的（接近 0 或 1），则分布将是偏斜的和细长峰状的（比正态更尖）。

(4) 方差的异质性　如果主效基因是分离的，则家系内方差将是异质性的，因为主效基因在一些家系中分离，但在其他家系中不分离。可是，这种和上一种方法几乎没有检测能力，因为检测与正态的离差或者检测同胞方差的异质性需要很大的样本；并且也会受到非分离的主效基因因素引起的偏离正态性或家系内方差异质性问题的困扰〔Hill 和 Knott (1990) 综述；Mayo，1989〕。Hammond 和 James (1970) 采用这些方法中的一些方法来检测分离群体中影响果蝇腹部刚毛数的主效基因，得出它们是无效的结论，特别是如果性状的遗传力低时。

(5) 亲子相似性　Karlin，Carmelli 和 Williams (1979) 以及 Karlin，Williams 和 Carmelli (1981) 提出一组不同的测验，集合起来称为“结构表现资料分析”(structured exploratory data analysis)。这种方法是以直觉显著原理 (intuitively obvious principle) 为基础，对于多基因遗传和没有主效基因而言，后裔的平均值更密切地相似于中亲值而不是单个亲本值，而如果主效基因是分离的，则相反的情形是真实的。这些测验中的一种，“主效基因指数应”用于已知群体内 *hg* 基因分离的小鼠家系中，这种测验成功地检测出主效基因 (Famula，1986)。可是，这些测验对分布的非正态性敏感，能导致错误的结论 (Mayo，Eckert 和 Nugrobo，1983；Kammerer，MacCluer 和 Bridges，1984)。

(6) 复合分离分析 (complex segregation analysis)　检测影响数量变异主效基因的最有力方法是复合分离分析 (Morton 和 MacLean，1974，由 Hill 和 Knott 评述，1990)，是专为人类的双亲和全同胞系谱而提出的。复合分离分析检测一个性状的遗传是用单个主效基因的分离来解释最好呢，还是用严格的多基因遗传或者是用主效基因加上微效多位点来解释最好。制定出一个完全混合的模型，它包括：等位基因频率和单主效基因位点的加性及显性效应，源于多重多基因因子的加性遗传效应，共同的环境效应和随机的环境效应。对一系列格外复杂的假设：纯环境模型、单基因模型、多基因模型和完全模型求参数的最大似然估值 (maximum likelihood estimates)。每个假设的显著性都通过比较数据的似然率来检验，即用似然比率检验 (likelihood ratio test) 法对得出的模型参数的最大似然估值与依假定最适零假设 (其检验参数为零时) 计算的似然估值作比较。似然函数很复杂，它的求值需要使用计算机，但方法通常能扩展至更为复杂的遗传假设中，包括基因多效性和与已知标记的连锁。复合分离分析已用于许多已知的人类疾病和疾病危险因子的遗传分析，在所有情况中都有显著的多基因组分，并且在某些情况下表明额外存在主效基因效应 (Sing 等评述，1988)。

通过上面任何一种方法可推断影响数量性状的主效基因存在，但仍然不能告诉我们这些基因是什么，对变异的多基因组分有贡献的位点仍然看不见。为此，目前分辨 QTL 的方法定向为识别所有的相关位点，这些位点可有一个效应范围 (从主效基因到实验的分辨限制)，将这些位点置于连锁图谱上，并且最终对相关的 DNA 序列进行分子克隆。

QTL 作图的方法

估计 QTL 效应和图谱位置的实验设计是单个基因作图标准方法的扩展，是以一个标记位点上的等位基因与连锁 QTL 上的等位基因之间连锁不平衡为基础的。因此 QTL 作图的要求是 (i) 一张多态标记位点的连锁图谱充分覆盖整个基因组，(ii) 群体或品系之内或之间数量性状存在变异。

标记位点

理想的标记位点应当是（i）高度多态的，以便成对的个体或品系在每一个位点上都可能携带不同的等位基因；（ii）丰富的，因此可以实现范围广泛的标记覆盖基因组；（iii）中性的，涉及有益数量性状和繁殖适合度两者；（iv）共显性的，因此在一个标记位点上所有可能的基因型都能被识别。这个最后要求不太严格，因为在某些设计中都能成功地使用显性/隐性标记，正如下面描述的那样。

直到最近，因为缺乏适合的标记而使作为数量变异基础的位点作图严重受阻。诸如血型抗原和电泳可辨别的等位酶基因等隐蔽的蛋白质变异往往可满足中性和共显性的标准，但对标记整个基因组来讲既不足够多态又不充分丰富。随着分子生物学的发展以及满足上述所有必需标准的 DNA 标记的发现，这种情况已经改变。第 1 章描述的限制性片段长度多态性（RFLP）、可变数目串联重复序列（VNTR）或小卫星位点、微卫星（或简单序列重复，SSR）位点都是适合的遗传标记。RAPD（DNA 随机扩增多态性）标记主要检测以任意 10 个碱基为引物经聚合酶链式反应（PCR）扩增后，产物所出现的 DNA 序列变异。10 个碱基序列无论何时侧面接触适当大小的基因组序列，标记 DNA 都可以扩增。RAPD 标记是显性的，因为 PCR 产物是典型地存在或不存在。RAPD 和 SSR 标记高度丰富和多态，它们的检测使用 PCR，意味着用它们构建连锁图谱比使用 RFLP 标记可能做的更快和更有效。在一些生物体中，例如果蝇和小鼠中，转座（位）因子或专宿反转录病毒序列（ecotropic retrovirus sequences）的基因组位置（插入位点）是高度变化的，它们在插入位点上的出现可用作一个显性多态标记。人、小鼠、大鼠、果蝇、很多家畜和植物物种的基因组的高密度分子标记连锁图目前是可得到的。

QTL 基因型

一般有两种方法用来识别 QTL 和作图：这些方法以有益性状不同的品系杂交为基础并且方法是以分离群体为基础的。定位 QTL 最有效的实验设计是采用 QTL 和标记位点上交替等位基因都固定的品系之间的杂交，因为 F_1 代的位点之间具有最大的连锁不平衡。更好的是，所有等位基因应当是联合的（association）；也就是说，增加性状值的等位基因在一个亲本品系中应当是纯合的，而降低性状值的等位基因则固定在另一个亲本品系中。如果亲本群体衍生于有益性状的趋异人工选择并且随后近交，那么联合 QTL 的纯合性和排列次序最可能得到满足。如果每一个品系都有一些增效的和一些减效的等位基因固定，则等位基因是分散的；在这些情况下，定位 QTL 仍然是可能的，不过与等位基因联合时相比，将发现较少的 QTL，其原因以后解释。

最普遍的情形是，亲本近交群体用来杂交以产生 F_1 代，然后 F_1 代要么与一个亲本品系，要么与两个亲本品系回交（BC 设计），或自群杂交以产生 F_2 代（F_2 设计）。在可忍受近交的物种内，F_2 能近交产生重组近交系（Oliverio，1979），这些重组近交系对作图也有用。对不能容忍近交和/或具有很长世代间隔的物种，品系杂交分析的原理仍然可以应用，条件是标记位点分离的品系是有用的，但这些品系的有益数量性状仍然是异化的，因而交替的 QTL 等位基因很可能被固定。在这种情形下，必须筛选亲本的标记构型（marker configurations）以在后代产生信息分离（即回交或 F_2 比率）。这是在人类遗传作图中使用的方

法。在随机交配群体中通过与标记位点的连锁检测 QTL 是最没有成效的，因为除非常紧密连锁的位点之外，所有其他位点似乎是连锁平衡的。

作为通过与标记位点连锁来识别 QTL 的基础原理，在概念上是简单的：根据标记位点上的基因型以及数量性状的表现型对个体评定。如果标记基因型类型之间的平均表型存在差异，那么我们能够推断与标记连锁的一个 QTL 存在。标记位点可以单个地或同时考察。通过与标记连锁检测出的 QTL 数目总是位点数的一个低估值，因为两个互相紧密连锁的 QTL 如果处于联合，则可能仅作为一个 QTL 出现，或者如果处于分散，则可能根本检测不到。在大多数实验中，很粗略地说，20 厘摩（cM）的图谱距离是分辨的限度，因此检测出的一个 QTL 是这种长度的一个染色体片段，它可能包含几个影响性状的位点，但不必是同一方向。因此当等位基因处于联合比当它们处于分散时，有可能检测出更多的 QTL。可是，借助更精确的方法，分辨的限度可以降低到大约 3cM。

单个标记分析

为了举例说明与单个标记有关的检测 QTL 的方法，考察一个标记位点（M）和一个 QTL（A），它们之间的重组频率为 c。假定一个亲本品系的基因型是 M_1A_1/M_1A_1，另一个亲本品系的基因型为 M_2A_2/M_2A_2。遵照图 7-1，A_1 和 A_2 纯合亲本的基因型值分别为 a 和 $-a$。F_1 个体（M_1A_1/M_2A_2）的基因型值是 d。F_1 亲本配子（M_1A_1 和 M_2A_2）每一个都以 $(1-c)/2$的频率产生，重组 F_1 配子（M_1A_2 和 M_2A_1）每一个都以 $c/2$ 的频率产生。F_1 随机交配产生 10 个可能的 F_2 基因型类型。每一个标记基因型类型对 F_2 平均值的贡献由每一个基因型频率乘以它的基因型值，然后在标记基因型类型内求和而获得。其过程示例在表 21-2。这里我们感兴趣的是标记类别之间的差异而不是标记类型对 F_2 平均值的贡献。因此我们需要实际平均值，这可通过 F_2 平均值的贡献除以那个标记类型的频率而获得，而这种频率是孟德尔分离比率，纯合子为 1/4，杂合子为 1/2。实际平均值给在表的最后一栏。以同样的方法计算回交中标记类型的平均值，但是它们更简单，因为只有 4 种基因型和两种标记类型。

从表 21-2 中我们可以看出，如果标记位点与 QTL 不连锁（即 $c=0.5$），那么所有标记类型有相同的期望平均值 $0.5d$。可是，如果 QTL 与标记位点连锁，那么标记类型平均值的下列比较是基因型值或加性效应 a 和显性离差 d 的函数：

$$(M_1/M_1 - M_2/M_2)/2 = a(1-2c) \qquad [21.1a]$$

以及在一些简化之后，

$$M_1/M_2 - [(M_1/M_1 + M_2/M_2)/2] = d(1-2c)^2 \qquad [21.1b]$$

因此，纯合标记基因型类型间一个数量性状平均值的显著差异可视为一个 QTL 与标记位点连锁的证据。可是，来自单个标记分析的 a 和 d/a 估计值因重组频率的混淆，一般是真值的低估计值达 $(1-2c)$。例如，纯合标记类型之间 0.2 个表型标准差的平均差异可归因子与标记完全连锁的具有这种效应的一个 QTL，或者归因于 46cM 远的效应为 $a=0.5$ 个标准差的一个 QTL（采用 Haldane 作图函数 $c=[1-exp(-2x)]/2$将以摩根表示的距离（x）与重组频率关联起来）。

表 21-2 具有一个标记位点（M）和一个连锁 QTL（A）的 F_2 中的基因型

F_2 基因型	频率	值	标记类型	频率	对 F_2 平均值的贡献	实际平均值
M_1A_1/M_1A_1	$(1-c)^2/4$	a	M_1/M_1	$\frac{1}{4}$	$a(1-2c)/4+dc(1-c)/2$	$a(1-2c)+2dc(1-c)$
M_1A_1/M_1A_2	$c(1-c)/2$	d				
M_1A_2/M_1A_2	$c^2/4$	$-a$				
M_1A_1/M_2A_1	$c(1-c)/2$	a	M_1/M_2	$\frac{1}{2}$	$d[(1-c)^2+c^2]/2$	$a[(1-c)^2+c^2]$
M_1A_1/M_2A_2	$(1-c)^2/2$	d				
M_1A_2/M_2A_1	$c^2/2$	d				
M_1A_2/M_2A_2	$c(1-c)/2$	$-a$				
M_2A_1/M_2A_1	$c^2/4$	a	M_2/M_2	$\frac{1}{4}$	$-a(1-2c)/4+dc(1-c)/2$	$-a(1-2c)+2dc(1-c)$
M_2A_1/M_2A_2	$c(1-c)/2$	d				
M_2A_2/M_2A_2	$(1-c)^2/4$	$-a$				

例 21.1

Sax(1923)报告了首例一个标记位点和一个数量性状之间的关联，即菜豆（*Phaseolus vulgaris*）一个色素位点和种子大小之间的关联。亲本品系之一是改良的 Yellow Eye 1317，其显性色素基因 *P* 纯合，种子平均重 48 厘克(cg)，另一亲本品系是 White 1228，其隐性色素基因 *p* 纯合，平均种子重 21.0 厘克。杂种 F_2 中色素位点上的基因型以及平均种子重量是：

基因型	*PP*	*Pp*	*pp*
种子重量(cg)	30.7	28.3	26.4

[F_2 色素基因型可从 F_3 中非色素(*pp*)后裔的存在或缺乏中推断出来。]等式[21.1*a*]给出 $a(1-2c)$的估计值为$(30.7-26.4)/2=2.15$cg。效应完全是加性的，因为标记杂合子几乎精确地处于两个纯合子的中间。等式[21.1*b*]给出$d(1-2c)^2=-0.25$cg。F_2 中 *PP* 和 *pp* 基因型之间种子重量的差异（4.3cg）解释了两个亲本品系之间种子重量总差异（27cg）的 16%。这是与标记关联的一种巨大效应，但关于一个 QTL 我们能得出的结论却是有限的。首先，部分或甚至全部效应可归因于对种子重量有基因多效性影响的 *P* 位点本身。其次，如果存在一个与 *P* 位点连锁的 QTL，只有它与标记没有重组时（$c=0$），它的效应（$2a$）才可正确地估计为 4.3cg。这个效应同样也可归因于有较大效应的一个不紧密连锁 QTL。第三，识别出的 QTL 可能不是一个而是两个或更多的连锁位点。

区间作图分析

通过考察被重组频率 c 分开的一对连锁标记位点（M 和 N），这两个标记位点对亲本品系中的交替等位基因是固定的，可以解决与一个具重组频率的标记位点连锁的 QTL 其加性和显性效应估计值受混淆的难题。在这种情形中，标记位点的图谱位置以及由此 c 是已知的。假定在两个标记位点之间存在一个 QTL（A），M 和 A 之间有重组频率 c_1，N 和 A 之间为 c_2；假定没有干扰，$c_1+c_2=c$。F_1 中标记位点之间的关系描绘在图 21-1 中。

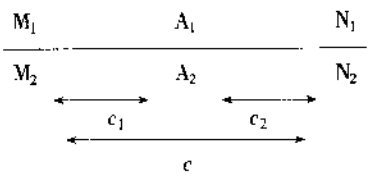


图 21-1　两个标记位点 M 和 N 与一个 QTL（A）之间的重组频率

这样，两个亲本品系的基因型是 $M_1A_1N_1/M_1A_1N_1$（基因型值为 a）和 $M_2A_2N_2/M_2A_2N_2$（基因型值为 $-a$），F_1 的基因型是 $M_1A_1N_1/M_2A_2N_2$（基因型值为 d）。在回交或 F_2 世代中可以进行分析。与纯合 $M_1A_1N_1$ 亲本回交的标记类型的期望平均值列在表 21-3 中，是用与表 21-2 相同的方式计算的。这里标记基因型的期望频率取决于两个标记之间的重组频率，因此具有 F_1 亲本配子的标记类型的平均值要除以 $(1-c)/2$，具有 F_1 重组配子的标记类型的平均值要除以 $c/2$。假定两个标记足够紧密连锁，于是双重组可以忽略。可给出 QTL 效应估计值以及它相对于侧翼标记的图谱位置的回交标记类型平均值之间的比较，可由等式［21.2］给出：

$$M_1N_1/M_1N_1 - M_1N_1/M_2N_2 = a - d \qquad [21.2a]$$

$$M_1N_1/M_1N_2 - M_1N_1/M_2N_1 = (a-d)(c_2-c_1)/c \qquad [21.2b]$$

回交设计的一个缺点是只有当 $d=0$ 时，QTL 加性效应的估计值才是无偏的，并且隐性或部分隐性 QTL 可能检测不到。这个问题可通过与两个亲本品系的回交或采用 F_2 设计来克服。在 F_2 中产生的 9 种标记类型的期望基因型值由 Haley 和 Knott（1992）给出。

表 21-3　具有两个连锁标记（M 和 N）以及一个连锁 QTL（A）的回交中的基因型。F_1 与 $M_1A_1N_1$ 进行回交

F_1 配子类型	频率	值	标记类型	频率	对回交平均值的贡献	实际平均值
$M_1A_1N_1$	$(1-c)/2$	a	M_1N_1/M_1N_1	$(1-c)/2$	$a(1-c)/2$	a
$M_1A_1N_2$	$c_2/2$	a	M_1N_1/M_1N_2	$c/2$	$(ac_2+dc_1)/2$	$(ac_2+dc_1)/c$
$M_1A_2N_2$	$c_1/2$	d				
$M_2A_1N_1$	$c_1/2$	a	M_1N_1/M_2N_1	$c/2$	$(ac_1+dc_2)/2$	$(ac_1+dc_2)/c$
$M_2A_2N_1$	$c_2/2$	d				
$M_2A_2N_2$	$(1-c)/2$	d	M_1N_1/M_2N_2	$(1-c)/2$	$d(1-c)/2$	d

例 21.2

在分子标记成为可用的之前，大多数 QTL 作图是用果蝇（*Drosophila*）刚毛数进行的，因为（1）具有异化值的选择品系是可用的，（2）覆盖在所有 3 条主要染色体上的主要突变体基因可用作标记，（3）可以每次研究一条染色体。早期的工作由 Thoday（1979）综述。下面的数据取自 Wolstenholme 和 Thoday（1963）。

使用的亲本染色体是第 3 号染色体，来自对高腹侧刚毛数选择的一个品系，测验种第 3 号染色体来自对刚毛数未加选择的品系，它有几个形态学标记。隐性基因 *clipped*（*cp*）和显性基因 *Stubble*（*Sb*），这两个标记为 12.9cM 的距离分开。为了确定在 *cp* 和 *Sb* 之间是否存在一个影响腹侧刚毛数的 QTL，具有选择染色体（HH）的果蝇与标记的测验种染色体品系的果蝇杂交。然后双杂合的 F_1 果蝇（*cpSb*/HH）与一个 *cp* + 测验种果蝇回交。回交杂种中 4 个标记基因型类型的平均刚毛数见表21-4。

检查清楚地表明存在一个与标记连锁的、影响刚毛数的 QTL；来自选择品系的 H 等位

基因与较高的数目相关联。从标记类型之间刚毛数的差异中，我们可以得到下面的估计值：

根据等式[21.2*a*]，$(a-d)=(1)-(4)=2.62$ 根刚毛，

根据等式[21.2*b*]，$(a-d)(c_2-c_1)/c=(2)-(3)=0.24$ 根刚毛，

表 21-4 不同基因型的刚毛数

基因型	刚毛数
(1) HH/*cp* +	20.62
(2) H*Sb*/*cp* +	19.19
(3) *cp*H/*cp* +	18.95
(4) *cpSb*/*cp* +	18.00

其中，c_1 是 QTL 和 *cp* 之间的重组频率，c_2 是 QTL 和 *Sb* 之间的重组频率。标记之间已知的距离给出 $c=0.129$。将 $(a-d)=2.62$ 和 $c=0.129$ 代入上面第二个等式，给出 $c_2-c_1=0.0118$。因此，从 $c_1+c_2=c$ 中我们能够估计出 $c_1=0.059$ 和 $c_2=0.070$。这样，QTL 被定位在两个标记 *cp* 和 *Sb* 之间，更接近于 *cp*。当只有一种回交时，我们不能区分 *a* 和 *d* 来估计 QTL 的效应和基因作用。

遗传和统计上的考虑

虽然 QTL 作图的原理易懂，但实践中关于统计分析的最优方法以及遗传解释出现很多问题。现在我们需要考虑在确定样本大小、标记的数目和密度、最优实验设计、适当的统计分析过程中所包括的问题以及有关实验资料遗传解释的局限性。

实验设计

我们对作图方法的描述在下述意义上过于简单化，即我们没有详细说明如何判断标记类型平均值（或两个标记类型平均值的一种比较，例如等式［21.1*a*］或［21.2*a*］）是显著不同的。目前假定为此目的应用 *t* 检验，我们想设计一种单个标记实验来检测 F_2 杂种中由与标记完全连锁（$c=0$）的一个 QTL 引起的两个纯合标记基因型之间的某一个任意值 δ 的差异。假定在每一个标记基因型类型中用于评定数量性状的个体数（*n*）足够大，以致于我们可以假设表型值是正态分布的，所需要的 *n* 由标准统计理论（Sokal 和 Rohlf，1981，p.263）给出为：

$$n\geqslant 2(z_\alpha+z_{2\beta})^2/(\delta/\sigma_W)^2 \qquad [21.3]$$

在这个表达式中，δ 是实验者想检测的作为显著的标记类型之间的最小差异，σ_W 是标记类型内的表型标准差；α 和 β 是设定的显著平准，α 是可接受的假阳性（false positives）误差率（Ⅰ型误差），β 是可接受的假阴性（false negatives）误差率（Ⅱ型误差）；*z* 是与其下标相对应的正态分布的纵坐标。例如，假定误差率设置为 $\alpha=0.05$ 和 $\beta=0.1$，给出 $z_\alpha=1.96$ 和 $z_{2\beta}=1.28$。于是，如果要检测标准化效应（δ/σ_W）为 1 或更大，那么等式［21.3］表明每一个标记类型需要的个体数是 $n=21$；可是，如果 δ/σ_W 被设定为 0.25，那么 $n=336$。因此检测具有中等或小效应的 QTL 需要很大的样本大小；对小样本，只有巨大效应的 QTL 才能被发现。类似的能力计算产生了下面关于 QTL 作图的不同实验设计相对效率的一般指导方针。

1. 区间作图比单个标记分析更好，因为对一单个标记，QTL 效应被 QTL 偏离标记的图谱距离所混淆了，正如我们已经看到的那样。区间作图也比单个标记分析更有效。对单个标记分析，检测一给定标准化效应所需要的样本大小增加了一个 $1/(1-2c)^2$ 部分

(Soller, Brody 和 Genizi, 1976)。

2. F_2 设计比回交设计更强有力。与一单个亲本的回交仅检测出杂合效应，对加性 QTL 的情形，这种杂合效应是 F_2 中检测出的纯合效应的一半，因此需要多达 4 倍的个体评定才能获得相同的能力。对于显性，一次回交将比 F_2 更有效，其余的不很有效（Soller, Brody 和 Genizi, 1976)。与两个亲本的回交对估计纯合效应是必要的，这也不及 F_2 设计有效。

3. 如果两个亲本群体对 QTL 上的交替等位基因是固定的，但对标记等位基因是分离的，则为获得与固定的标记等位基因相同的能力，样本大小要增加。例如，对单个标记分析来说，增加了一个 $1/(q_1-q_2)^2$ 因子，其中 q_1 和 q_2 是两个品系中标记的基因频率 (Soller, Brody 和 Genizi, 1976)。甚至对 $q_1=0.8$ 和 $q_2=0.3$ 这样一种有利的情形（即标记等位基因频率是相对极端的)，样本的大小增加 4 倍将是必要的，假定所有 F_2 个体都被评定。可是，如果只用产生信息后代的那些 F_1 交配，那么可以获得大量节约。信息交配是标记为杂合的两个个体之间的交配以及一个标记为杂合子和一个标记为纯合子的个体之间的那些交配。需要多少个个体的细节参见 Beckmann 和 Soller (1988)。

4. 检测两个标记基因型之间平均值差异的能力不取决于差异的绝对值（δ)，而是依赖于用标记类型内的标准差度量的差异（δ/σ_W)。因此降低 σ_W 的策略能够产生较高的检测能力。降低 σ_W 需要更准确的表型值估计值。这可以通过对分离世代 F_2 或回交的个体进行后裔测验或通过近交产生重组近交系来达到 (Thoday, 1961; Lander 和 Botstein, 1989; Soller 和 Beckmann, 1990)。形成重组近交系有缺点，就是 QTL 和标记之间的连锁为随后世代中进一步的重组所降低。可是，在果蝇（*Drosophila*）中，在没有进一步重组的少数世代中便能合成重组的等基因系，因此对这个物种，这种方法是最好的 (Long 等, 1995)。

多重测验

上面的讨论涉及的是与一单个标记或一对标记连锁的 QTL。实际上，在整个基因组上分布很多标记，要依次地测验每一个（或每一对）标记与 QTL 的连锁。而且，典型使用的亲本品系在多个性状上都是歧化的，所有这些性状在分离世代中都要评定，因此与标记连锁的测验对每一个性状都要重复进行。这意味着一些“显著”的正联合将偶然发生，必须设定一个更严格的显著平准。假阳性的数目随着测验的次数而快速增加。假定一组 n 个独立测验的期望显著平准为 $\alpha=0.05$，那么每一个独立测验的平准必须设定为 α/n。例如，如果对 50 个独立标记测验单个标记联合，那么每一个测验的显著平准应当是 $\alpha=0.001$。用 $z_{0.001}=3.291$ 对等式 [21.3] 的重新评估表明样本大小必须加倍才能获得相同的能力。可是，测验是不独立的，因为在一些标记之间存在连锁，因而一些测验将是部分多余的。这样，将测验作为独立的来处理设定了太严格的显著平准。当采用最大似然法来分析数据时，正如下面解释的那样，对标记的非独立性可以作适当的考虑。

最大似然估计

为示例目的在有关 t 测验的行文中讨论了不同设计的相对效率。可是，这种测验的使

用在技术上是不合适的，因为它是以假定标记类型基因型内表型的正态分布为基础的。对表 21-1 的审查表明，这不是真实的：F_2 纯合标记基因型的每一个都包含 3 种 QTL 基因型，因此分布是 3 个正态分布的混合（进一步假定环境方差是正态分布的）。除了极端偏离正态性之外，诸如极端偏离是由区分两个亲本品系的极少量巨效 QTL 所引起的情况之外，用 t 测验（或方差分析）检测 QTL 和标记的连锁都是最有活力的。但是，基于最大似然的较复杂高级的方法对参数估计和显著性检验更适合，因为它们考虑了与分离 QTL 有关的标记基因型的正确分布特性。详细细节参见 Lander 和 Botstein（1989）以及 Knott 和 Haley（1992）。

这种方法概括地介绍如下。根据观察资料（每一个标记类型中个体数目和它们的表型）以及要估计的参数（c 以及 QTL 基因型的平均值和方差）来指定一种似然函数。给未知参数设计试验值，并为每一个试验值用迭代计算机程序寻找似然函数 L。使 L 最大的试验值就是未知参数的最大似然估计值。因此最大似然估计对单个标记情形是有利的，因为标记和 QTL 之间的重组频率在模型中得以考虑并被估计出来。显著性检验是比率 L/L_0（它服从 χ^2 分布）的对数（底数为 10），其中 L 是观察的最大似然值，L_0 是对没有 QTL 分离的零假设计算出的似然值。

遵从人类连锁作图的惯例，在区间作图方面，似然比率的对数称为 LOD 分数（对于“优势对数”）。详细细节见 Lander 和 Botstein（1989）。对区间内 QTL 变化的位置计算 LOD 分数；c 和 a 的最大似然估计值就是 LOD 分数最大时的数值。估计值经常以图解的形式出现，是一张以 LOD 分数对染色体位置（cM 距离）作的图。图 21-2 展示了一个实例。高于按 5% 总误差率判定区间中 QTL 存在为显著的 LOD 分数值，必须从独立的区间检测数中计算，这依次取决于标记密度、染色体数和基因组大小。对这些参数假定值广泛范围的典型 LOD 阈值在 2 和 3 之间，对每一个检测区间这大致相当于 $\alpha = 0.001$（Lander 和 Botstein，1989）。在图 21-2 展示的玉米籽粒产量的研究中，有 8 个连锁标记，它们的位置展示在 X 轴上，它们之间有以 cM 表示的区间。用三角形标记的位置是使 LOD 分数最大的 QTL 的位置，对一个 QTL 来说，这是最可能的位置。LOD 分数为 2 的水平线表示整个实验显著性（$\alpha = 0.05$）的阈值。存在一个影响产量的 QTL 得到了明显的证实。

多个 QTL

很显然，选择用于作图的亲本品系很可能在影响有益性状的很多位点上有不同。迄今我们忽略了与所考察的标记不连锁的 QTL 的存在，但在 F_2 或回交世代中它们仍然在分离。因此标记基因型类型内的方差 σ^2_W 将包含由不连锁 QTL 分离引起的一种遗传组分。而且，通过与标记位点连锁识别的 QTL 不是平常遗传意义上的位点（即，包括在产生的多肽链中的 DNA 片段），而是有效因子（第 12 章），可能包含影响性状的几个位点。因此，当测验揭示出一个与标记连锁的 QTL 时，观察到的效应可能是由两个或更多个位点，而不仅仅是由一个位点引起的。于是观察到的效应是两个或更多个位点效应的总和。连锁位点上的等位基因可能处于联合（所有基因增加性状值）或处于分散。因此，与观察到的总和效应相比，如果处于联合，则单个位点的效应可能较小，或者如果处于分散，则较大。

有同时处理多个 QTL 的方法，这会改进图谱位置和效应的估计值。这些方法的描述

见 Jansen 和 Stam（1994）以及 Zeng（1994）。考虑多个 QTL 的方法对实验资料的分析是最理想的。

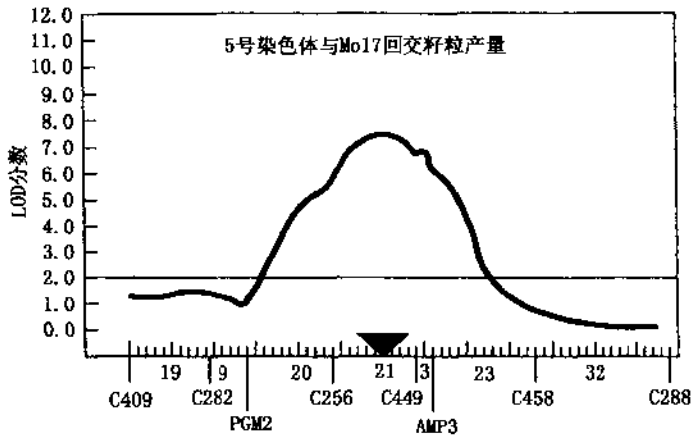


图 21-2　用优良玉米自交系 B73 和 Mo17 的 F_1 与 Mo17 亲本的回交所做玉米第 5 号染色体上 LOD 分数对影响籽粒产量 QTL 的染色体位置图。同功酶（PGM2 和 AMP3）以及 RFLP（标明 C—）标记的位置绘在 X 轴上，标记之间的距离以 cM 计。一个产量 QTL 的最可能位置用全黑三角形标明，对应于 LOD 分数的最高值（C.W.Stuber 热情地提供未发表的图和数据）

实验结果

对很多生物体许多性状的实验研究充分证实的一般结果是，有多个 QTL 分散在整个基因组中。因此，数量变异的多基因模型得以清楚地确认。而且，“无穷小”模型，即有几乎无穷数目的位点，每一位点具有不可觉察的小效应的模型显然是错误的。实验结果概述如下。

位点数目

所有 QTL 数目的估计值都是影响性状位点真实数目的最小估计值，有 3 个理由：（i）实验在分离紧密连锁位点的能力方面是有限的，（ii）必定总是存在其他位点，它们效应太小以致于特定大小的试验不能检测到，（iii）找到的位点是区分被比较的两个品系的那些位点；其他位点可能在其他品系中找到。

影响黑腹果蝇（*D. melanogaster*）刚毛数的 QTL 数目的一些例子给在表 21-5 中。有 3 条较大的染色体和 1 条很小的染色体；大多数实验仅在 1 条或两条染色体上寻找 QTL。测验的成对亲本染色体要么来自选择高刚毛数的一个品系和一个未选择的测交群（Wolstenholme 和 Thoday，1963；Spickett 和 Thoday，1966），要么来自选择高和低刚毛数的品系（Shrimpton 和 Robertson，1988*a*，*b*；Long 等，1995）。找到的 QTL 解释了测验的亲本染色体之间刚毛数的全部或大部分差异，这意味着在测验的染色体上所有重要的 QTL 都检测到了。

最近，番茄、玉米、小鼠和其他物种的分子标记图谱已用来对影响各种性状的 QTL

作图。一些例子给在表 21-6 中；更全面的综述见 Tanksley（1993）。检测出的 QTL 解释了 F_2 或 BC 世代中的大部分表型方差。因为不知道用于作图的群体中性状的遗传力，所以我们不能说解释了多少遗传方差，但大多数重要的 QTL 一定被检测出来了。

表 21-5　影响黑腹果蝇（*D. melanogaster*）腹侧（ST）和腹部（AB）刚毛数的 QTL 数目。Long 等（1995）用转座（位）因子插入位点作标记；在其他研究中，采用形态突变体作标记

性状	测验的染色体	标记数目	QTL 数目	亲本差异%	参考文献
ST	3	4	2	114	Wolstenholme 和 Thoday(1963)
ST	1,2,3	10	5	87.5	Spickett 和 Thoday(1966)
ST	3	7	18	103	Shrimpton 和 Robertson(1988*b*)
AB	1,3	45	7	89.9	Long 等(1995)

表 21-6　用与分子标记连锁对 QTL 作图的实例。表型方差是 F_2 或 BC 群体的

物种	标记数目	性状	QTL 数目	表型方差%	参考文献
番茄	70	果体	6	58	Paterson 等(1988)
(*Lycopersicon* sp.)		可溶物	4	44	
		果实 pH	5	48	
	71	果体	7,9	72	Paterson 等(1991)(1)
		可溶物	4,3,5	44	
		果实 pH	5,8,2	34	
玉米	76	籽粒产量	6,8	61,59	Stuber 等(1992)(2)
(*Zea mays*)		株高	3,5		
		果穗数	2,2		
人类	289	Ⅰ型糖尿病	18		Davies 等(1994)

(1) 估计值来自在两种或三种环境中生长的植株。表型方差%是针对列出的第一个值的。

(2) 估计值来自与每一个亲本品系的回交，每一个值是对 6 种环境平均得来的。

基因效应

加性效应　在大多数研究中，已发现具有巨大加性效应的 QTL。可是，并非所有的 QTL 都有相等的效应，像果蝇（*Drosophila*）刚毛数以及玉米生长和繁殖性状那样，不同性状所呈现的一般模式是：具有巨大效应的几个 QTL 解释了亲本品系之间的大部分差异，较大数目的微效 QTL 解释了差异的剩余部分。图 21-3 展示了影响果蝇（*Drosophila*）腹侧刚毛数的 QTL 效应的分布（Shrimpton 和 Robertson，1988*b*）。在第 3 号染色体上发现了 18 个 QTL，它们的效应（a）从 0.4 σ_P 变化到 1.9 σ_P。只有 3 个的效应比 1.0 σ_P 大。少量效应低于 0.6 σ_P 的 QTL 并不意味着真实的效应分布落在较低的效应水平上。面是较小效应的那些 QTL 的检出概率较低，因此它们中只有一些出现在数据中。或许真实分布偏斜的越甚便有更多的微效 QTL。

显性度　QTL 的显性度跨越从加性到完全显性甚至到超显性的整个范围。这例释在图

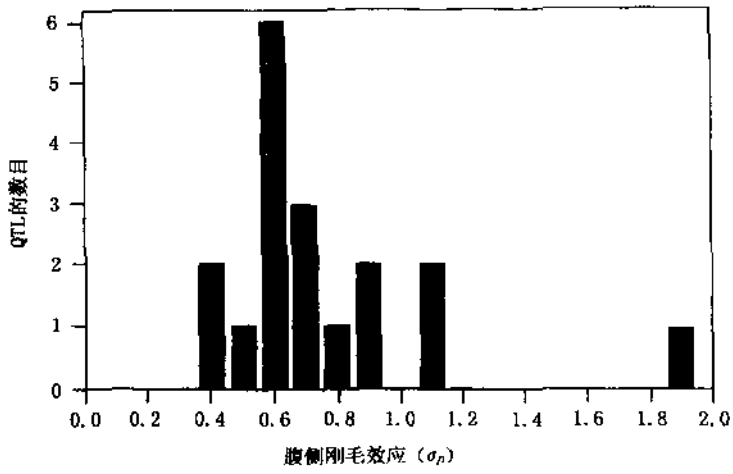


图 21-3　黑腹果蝇（*D. melanogaster*）第 3 号染色体上影响腹侧刚毛数的 QTL 效应（a）的分布，以表型标准差为单位

（引自 Shrimpton 和 Robertson，1988*b*）

21-4 中，该图展示出在番茄种间杂交 F_2 中识别出的 74 个 QTL 之间的显性度（d/a）分布。增效和减效等位基因的显性是同等频繁的。难于精确估计比率 d/a，因为它同时受 a 和 d 抽样误差的控制。取极端值的估计值特别容易受到怀疑，因为极端值是由极低的 a 值产生，于是比率对 a 的抽样误差非常敏感。超显性的情形尤其需要进一步确证。

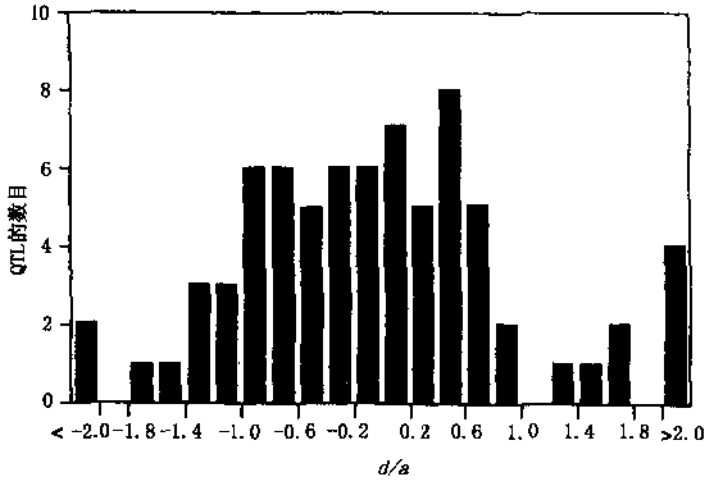


图 21-4　影响番茄营养性状的 74 个 QTL 间显性度（d/a）的分布，它们是在栽培番茄和野生种之间杂交产生的 F_2 中检测出的

（引自 deVicente 和 Tanksley，1993）

观察到的 QTL 显性度范围与具有巨大表型效应的突变体正常表现出来的完全显性或

隐性进行对比。可推测的是，对数量变异负责的 QTL 等位基因实际上在位点生产的蛋白质中产生数量差异，而不是总体缺乏。这应当是说，如果基因产物是一种酶，那么改变的是活性，或者，如果基因产物是一种转录因子，那么改变的是结合效率。但是分离中几乎没有酶活动。相反，它们是产生最终产物的代谢途径中的步骤，正是这种终产物产生的数量或速度影响我们所见的表型。对代谢途径中酶动力学的考察导致下列结论：两个等位基因之间酶活性存在巨大差异时，具有较高活性的等位基因将倾向于是显性的，但存在小差异时，等位基因将倾向于以加性方式起作用（Kacser 和 Burns，1981；还可参见 Hartl，Dykhuizen 和 Dean，1985；Dean，Dykhuizen 和 Hartl，1988）。

上位 QTL 之间的互作难于检测，因为需要很大数目的实验。已在影响果蝇（*Drosophila*）刚毛数的 QTL 之间观察到了强烈的上位互作（Spickett 和 Thoday，1966；Shrimpton 和 Robertson，1988*a*；Long 等，1995）。Long 等（1995）检测到的上位互作效应大约与主要效应同样大小。虽然上位可能是普遍的，但这并不一定意味着上位对随机交配群体的遗传方差有大比例的贡献。

相关效应 鉴于性状之间遗传相关的普遍发生，发现很多 QTL 影响一个以上的性状就不惊奇了。例如，果蝇（*Drosophila*）的几个实验（例如，Shrimpton 和 Robertson，1988*b*；Long 等，1995）已经发现，对测验中的选择亲本品系的刚毛性状具有最大影响的 QTL，也影响其他的刚毛性状。在玉米的 82 个被测性状中，发现多达 78 个性状受一个单个染色体区域影响（Edwards，Stuber 和 Wendel，1987）。

一些实验测验了不同环境中的同一杂种（Paterson 等，1991；Stuber 等，1992）。正如所预料的那样，如果不同环境中的表现在遗传上是相关的，那么某些 QTL 便在一种以上的环境中检出。

一致性

如果被测品系偶然固定了效应相似的等位基因，那么某些“显著”的 QTL 可能错误定位并且对群体内和群体间显著变异负责的 QTL 也会漏检。因此，应当用相同的和不同的品系经重复试验来确认 QTL。采用从不同基础群体中获得的品系，已对影响果蝇刚毛数的 QTL 作了图（Breese 和 Mather，1957；Thompson 和 Thoday，1974；Shrimpton 和 Robertson，1988*b*；Long 等，1995）；对影响番茄果实性状（Weller，Soller 和 Brody，1988；Paterson 等，1988，1991）和玉米花序性状（Doebley 和 Stec，1991，1993）的 QTL 也作了图。当不同品系之间的一致性相差甚远时，对估计的 QTL 图位置和效应，有聚集在同一染色体区域的明显趋势。可能，物种之间的比较更有趣。用一种可比的 RFLP 连锁图谱来对影响豇豆（*Vigna unguiculata*）和绿豆（*Vigna radiata*）种子重量的 QTL 独立交叉作图，结果表明对两个物种种子重量有最大效应的 QTL 作图到相同的位置上（Fatokun 等，1992）。在数目有限的共同位点上的等位变异负责群体内和群体间，甚至物种间的变异都是可能的。因此，模型生物和更有实践价值的物种的 QTL 之间可能有的同源性充分证明，诸如人类疾病或家畜生产性状的小鼠模型等实验上易控制的模型系统的采用是十分正确的。

从 QTL 到基因

典型的 QTL 作图试验成功地将 QTL 定位到大约 20cM 区域内，这其中暗藏着许多影响

性状的位点。一个区间内的连锁 QTL 通过进一步的重组能分开，并且如果区间内没有标记是有用的，它们的图位置可经后裔测验确定在大致 3cM 范围内（Thoday，1961；Lander 和 Botstein，1989；还可参见 McMillan 和 Robertson，1974），或者如果区间包含标记，则可通过精细尺度作图来实现这一点（Paterson 等，1990）。这种水平的分辨率对于在选择性育种方案中使用 QTL 是足够的，但是，对于识别单个位点上负责数量性状表型差异的等位基因差异仍然差若干数量级。识别影响数量性状的实际位点，对人类多基因疾病的风险评估、对转基因技术应用于农业上重要的性状以及对根据等位基因频率和效应来描述数量变异的遗传基础都是必要的。有两种方法可用来识别在基因组一特定区域作为一个 QTL 而识别的一个基因；它们是位置克隆以及数量性状表型变异与在同一区域的“候选”位点上的多态标记的联合。

位置克隆要求有关位点的图谱位置已知在 0.3cM 区间内，这大约是目前获得的克隆载体中所包含的基因组插入部分的大小。这可通过实验生物中高分辨率的减数分裂图谱来达到，或通过对随机交配群体筛选位于已作图 QTL 区域上的多态标记来达到，而且这些多态标记与数量性状表型处于强烈的连锁不平衡之中。从这一点来看，识别有益基因和确定什么是与不同效应等位基因相关联的多态性，在概念上是易懂的，但技术上是极其困难的。这种方法对于由效应巨大的突变等位基因所识别的位点是可行的〔由 Takahashi，Pinto 和 Vituterna（1994）所综述〕，并已经用来识别影响人类疾病的单个位点〔例如，囊性纤维变性和亨廷顿氏疾病；由 McKusick（1994）所综述〕。对微效应的位点，困难将显著加剧，方法还没有用来分辨 QTL 到单个位点的水平。

从定位区域通往基因的最普遍策略是候选基因方法。很多已知功能的位点经常从未知位点图谱的区域中被识别出来并被克隆。能潜在地产生与未知位点的表型相关联的已知位点是候选位点，方法就是要寻找这个区域内每一个候选位点上表型与分子多态性的关联。一个 QTL 的近似位置不是提出在功能上与一个性状相关联的已知候选位点的先决条件。这种方法成功地应用于人类疾病的实例是脱脂蛋白 E（ApoE）等位基因与血清总胆固醇和心脏病的关系〔Sing 等（1988）综述〕；ApoE－4 等位基因与迟发作的 Alzheimer 疾病的关系（Corder 等，1993），以及 I 型糖尿病与主要组织相容性 HLA 区域和胰岛素基因区域变异的关系〔Davies 等（1994）综述〕。

一般而言，如果我们要将候选位点方法应用于数量性状，那么我们需要能够提出在功能上与性状相关的有关候选位点。对上面描述的人类疾病，患病和未患病个体之间生物学差异的详细了解指导了候选位点的选择。对其他的数量性状，潜在的候选位点可能包含在引起有益表型的生物化学和发育途径中的那些位点。例如，可能解释奶牛的牛奶成分变异的候选位点是酪蛋白和乳球蛋白位点（Pirchner，1988）；果蝇刚毛数数量变异的候选位点可能是对刚毛发育必需的那些位点（Mackay 和 Langley，1990）。因此，数量性状的很多候选位点已借助具有巨大突变效应的等位基因功效而识别出来。将候选位点方法应用于 QTL 是以下列假设为基础的：在这些位点上具有的微效应（即不够大从而不能描述为孟德尔突变）的“同等位基因”的分离引起了数量变异（Thompson，1975；Mackay，1985*b*；Robertson，1985）。

例释这种方法能力的实例是牛 β－*Lg* 位点与牛奶中乳清蛋白质浓度的大约一半遗传方差相关联的证实（Pirchner，1988）以及果蝇 *achaete－scute* 和 *scabrous* 基因区域每一个分

别解释 X 和第 2 号染色体超过 20% 的腹部和腹侧刚毛数遗传变异的证实(Mackay 和 Langley，1990；Lai 等，1994)。而且，由 QTL 作图识别的基因组区域的图谱位置经常大致地与影响度量性状的一个已知主效基因的位置一致。影响果蝇刚毛数的 QTL 近似地定位到主要刚毛位点 *achaete - scute*，*Notch*，*bobbed*，*daughterless*，*scabrous*，*extramacrochaetae*，*hairy*，*malformed abdomen*，*Delta* 和 *Enhancer of split* 的位置上（Shrimpton 和 Robertson，1988*b*；Long 等，1995)。一个矮化位点 *compact* 定位在与对玉米多个产量性状具有巨大影响相关联的区域上（Edwards，Stuber 和 Wendel，1987)，玉米 *teosinte - branched* 位点作图在对两个玉米—大刍草杂种花序性状有巨大影响的区域上（Doebley 和 Stec，1991，1993)。

根据复杂遗传学而不是统计描述来理解数量性状的前景是光明的。各个基因组项目正在产生人类、模型生物、农业上重要物种的高密度连锁图，这些连锁图经常表现出分类单位之间连锁群的显著保守性。随着检测与分子标记相联系的分离 QTL 的实验杂交和系谱分析的改良统计方法的发展，以及随着在单个基因水平上分辨 QTL 的潜力，对数量变异的孟德尔遗传学基础的描述是能够获得的。

习 题

21.1 下面的数据来自一个实验（Long 等，1995)，其中影响黑腹果蝇刚毛数的 QTL 作图在 3 条巨大染色体的两条上。这里的数据仅涉及第 3 号染色体。对腹部刚毛的双向选择应用于一个果蝇群体，它衍生于一个自然群体的大样本。在选择 25 代之后，高和低品系的每一个其所有染色体都纯合了，但很小的 4 号染色体除外，它被忽略了（这可通过操作果蝇染色体的特殊方法来做到)。然后，为了减少不想要的变异，以及具有一个共同的背景来比较第 3 号染色体的效应，高品系的第 1 和第 2 号染色体被它们的低品系同源染色体所代替。于是这些纯合的品系仅在它们第 3 号染色体上有所不同；它们是用于分析的“亲本品系”。亲本品系进行杂交。借助特殊的方法，来自已经发生重组的 F_1 雌性第 3 号染色体形成纯合，而第 1 和第 2 号染色体的低品系同源染色体作为纯合子保留。然后，重组的第 3 号染色体在等基因系中增殖，在每一个等基因系中，一个纯合的重组染色体可在很多个体中复制。在每一个等基因系中对 40 个个体记录了腹部刚毛数以及腹侧刚毛数。

	腹部刚毛		腹侧刚毛	
亲本品系	高	低	高	低
平均值	21.4	7.7	20.9	16.3
区间 1	HH	LL	HH	LL
n	13	31	13	31
平均值	14.3	11.9	19.5	16.7
平方和（校正）	277.2	374.6	17.4	11.2
方差	23.1	12.5	1.4	0.4

用来定位 QTL 的标记是称为 *roo* 的一种转座（位）因子的插入位点。在不同亲本品系的第 3 号染色体上有 29 个位点，与邻近标记之间的平均距离为 3.8cM。这里的数据涉及区间之一，即位于染色体一端的区间 1。在这个区间中，只考虑了没有重组的品系。这样有两个标记类型 HH 和 LL 纯合子，它们中分别有 13 和 31 个品系。除这些之外，还有 5 个品系在这个区间中发生了重组。

关于这个区间的 QTL，能得出什么结论？标记类型内方差的来源是什么？腹侧刚毛比腹部刚毛具有低得多的方差，其原因可能是什么？

[解答序号：141]

附 录 表

附录表 A 截断的正态分布——大样本。p = 超过截断点 T 值的群体比例。$x = T$ 值距平均值的离差，以标准差为单位。i = 个体值大于 T 值的个体平均离差，以距群体平均值的标准差为单位。对大于 50% 的 p 值：按（$1-p$）取 x 和 i 表值，给 x 一个负号；将 i 乘以 $(1-p)/p$，保留正号。来自 p 的线性内插值的误差是正的，当 $p>0.10\%$ 时，在 x 和 i 中误差最大，大约都是 +0.001（节选自 Falconer，1965a）。

p%	x	i	p%	x	i	p%	x	i
0.01	3.719	3.960	0.75	2.432	2.761	10	1.282	1.755
0.02	3.540	3.790	0.80	2.409	2.740	11	1.227	1.709
0.03	3.432	3.687	0.85	2.387	2.720	12	1.175	1.667
0.04	3.353	3.613	0.90	2.366	2.701	13	1.126	1.627
0.05	3.291	3.554	0.95	2.346	2.683	14	1.080	1.590
0.06	3.239	3.507	1.00	2.326	2.665	15	1.036	1 554
0.07	3.195	3.464				16	0.994	1.521
0.08	3.156	3.429	1.0	2.326	2.665	17	0.954	1.489
0.09	3.121	3.397	1.2	2.257	2.603	18	0.915	1.458
0.10	3.090	3.367	1.4	2.197	2.549	19	0.878	1.428
			1.6	2.144	2.502	20	0.842	1 400
			1.8	2.097	2.459	21	0.806	1.372
0.10	3.090	3.367	2.0	2.054	2.421	22	0.772	1.346
0.12	3.036	3.317	2.2	2.014	2.386	23	0.739	1.320
0.14	2.989	3.273	2.4	1.977	2.353	24	0.706	1.295
0.16	2.948	3.234	2.6	1.943	2.323	25	0.674	1.271
0.18	2.911	3.201	2.8	1.911	2.295	26	0.643	1.248
0.20	2.878	3.170	3.0	1.881	2.268	27	0.613	1.225
0.22	2.848	3.142	3.2	1.852	2.243	28	0.583	1.202
0.24	2.820	3.117	3.4	1.825	2.219	29	0.553	1.180
0.26	2.794	3.093	3.6	1.799	2.197	30	0.524	1.159
0.28	2.770	3.070	3.8	1.774	2.175	31	0.496	1.138
0.30	2.748	3.050	4.0	1.751	2.154	32	0.468	1.118
0.32	2.727	3.030	4.2	1.728	2.135	33	0.440	1.097
0.34	2.706	3.012	4.4	1.706	2.116	34	0.412	1.078
0.36	2.687	2.994	4.6	1.685	2.097	35	0.385	1.058
0.38	2.669	2.978	4.8	1.665	2.080	36	0.358	1.039
0.40	2.652	2.962	5.0	1.645	2.063	37	0.332	1.020
0.42	2.636	2.947				38	0.305	1.002
0.44	2.620	2.932				39	0.279	0.984
0.46	2.605	2.918	5.0	1.645	2.063	40	0.253	0.966
0.48	2.590	2.905	5.5	1.598	2.023	41	0.228	0.948
0.50	2.576	2.892	6.0	1.555	1.985	42	0.202	0.931
			6.5	1.514	1.951	43	0.176	0.913
			7.0	1.476	1.918	44	0.151	0.896
0.50	2.576	2.892	7.5	1.440	1.887	45	0.126	0.880
0.55	2.543	2.862	8.0	1.405	1.858	46	0.100	0.863
0.60	2.512	2.834	8.5	1.372	1.831	47	0.075	0.846
0.65	2.484	2.808	9.0	1.341	1.804	48	0.050	0.830
0.70	2.457	2.784	9.5	1.311	1.779	49	0.025	0.814
0.75	2.432	2.761	10.0	1.282	1.755	50	0.000	0.798

附录表 B 截断的正态分布——小样本。当 n 个个体从总数为 N 的群体中选择而来时，表值是选择强度 i。N 的线性内插值的误差是负的，最大的误差大约是 -0.0075；n 的内插值给出正误差，最大值大约是 $+0.006$。（节选自 Becker，1984，其中可以找到极其广泛的表。）

n	N									n
	2	3	4	5	6	7	8	10	12	
1	0.564	0.846	1.029	1.163	1.267	1.352	1.424	1.539	1.629	1
2	—	0.423	0.663	0.829	0.954	1.055	1.138	1.270	1.372	2
3	—	—	0.343	0.553	0.704	0.821	0.916	1.065	1.179	3
4	—	—	—	0.291	0.477	0.616	0.725	0.893	1.019	4
5	—	—	—	—	0.253	0.422	0.550	0.739	0.877	5
6	—	—	—	—	—	0.225	0.379	0.595	0.748	6
7	—	—	—	—	—	—	0.203	0.457	0.627	7
8	—	—	—	—	—	—	—	0.318	0.509	8
9	—	—	—	—	—	—	—	0.171	0.393	9
10	—	—	—	—	—	—	—	—	0.274	10

n	N									n
	14	16	18	20	25	30	40	50	60	
1	1.703	1.766	1.820	1.867	1.965	2.043	2.161	2.249	2.319	1
2	1.456	1.525	1.585	1.638	1.745	1.829	1.957	2.052	2.127	2
3	1.271	1.347	1.412	1.469	1.584	1.674	1.810	1.911	1.990	3
4	1.119	1.201	1.271	1.332	1.455	1.550	1.694	1.799	1.882	4
5	0.986	1.075	1.150	1.214	1.345	1.446	1.596	1.705	1.792	5
6	0.866	0.962	1.042	1.110	1.248	1.354	1.510	1.624	1.713	6
7	0.755	0.858	0.943	1.016	1.161	1.271	1.434	1.552	1.644	7
8	0.650	0.760	0.851	0.928	1.081	1.196	1.365	1.487	1.582	8
10	0.447	0.577	0.681	0.767	0.936	1.061	1.242	1.372	1.472	10
15	—	0.118	0.282	0.405	0.624	0.777	0.991	1.139	1.252	15
20	—	—	—	—	0.336	0.530	0.782	0.951	1.076	20

n	N									n
	70	80	100	150	200	250	300	350	400	
1	2.377	2.427	2.508	2.649	2.746	2.819	2.878	2.927	2.968	1
2	2.189	2.242	2.328	2.478	2.580	2.657	2.718	2.769	2.813	2
3	2.055	2.111	2.201	2.357	2.463	2.543	2.607	2.660	2.705	3
4	1.950	2.008	2.101	2.263	2.372	2.455	2.520	2.574	2.621	4
5	1.862	1.922	2.018	2.185	2.297	2.382	2.449	2.504	2.552	5
6	1.786	1.848	1.947	2.118	2.233	2.320	2.388	2.445	2.493	6
8	1.659	1.724	1.828	2.007	2.127	2.217	2.288	2.346	2.396	8
10	1.553	1.621	1.730	1.916	2.040	2.132	2.206	2.266	2.317	10
15	1.342	1.417	1.536	1.738	1.871	1.970	2.048	2.112	2.166	15
20	1.175	1.257	1.386	1.601	1.742	1.846	1.928	1.995	2.051	20
25	1.032	1.121	1.259	1.488	1.636	1.745	1.830	1.900	1.958	25

符号集注

方括号中的数字指的是采用该含义的那一章的序号。
这里没有列出限制使用的一些含义。

A_1, A_2	所考察的一个位点上的两个等位基因
A	育种值
a	纯合体 A_1A_1 的基因型值,作为距纯合体中值的离差
B 或 b	作为下标,表示家系间或组间
b	回归系数;例如 b_{OP} = 亲子回归
c	[1,21]重组频率
CR	相关选择反应
D	显性离差
d	杂合体的基因型值,作为距纯合体中值的离差
E	环境离差
e^2	$=1-h^2$
E_c	共同环境;即家系平均值距群体平均值的环境离差
E_g	永久或一般的环境效应
E_s	暂时或特殊的环境效应
E_w	家系内环境;即个体距家系平均值的环境离差
F	近交系数
F_1	系群或群体间杂交的第一代
F_2	F_1 间随机交配所产生的杂交种的第二代
FS	全同胞
f	共祖率 = 血缘系数
f	用于下标,指雌性
f	[13]用于下标,指家系间
G	基因型值

GCA	一般配合力
H	杂合基因型 A_1A_2 的频率
H	[14]杂种优势的大小;即杂种平均数距中亲值的离差
HS	半同胞
h^2	遗传力('狭义')
I	由上位引起的互作离差
I	[13,19]选择指数
i	选择强度;即以表型标准差为单位的选择差
j	[20]稳定选择强度的一种度量
k	各种含义的数字。在[4,10,20]中指家系大小,即家系中的个体数
L	[2]负荷
L	[4,11]世代长度
M	群体平均值
m	[18]群体平均值
m	下标,指雄性
m	[10]配偶育种值之间的相关
N	群体大小;即一个群体或系群中繁殖个体的数目
N	[10,13]家系数
N_e	有效群体大小
n	各种含义的数字。在[10,13]中特指每个家系的后代数
O	后裔
P	亲本
$\bar{P}$	中亲值
P	纯合基因型 A_1A_1 的频率
P	随机交配指数(泛交指数) = 1 − *F*
P	表型值
p	对性状具增效作用的等位基因 A_1 的基因频率
p	[11,18]中选比例或超过正态分布截断点的比例
pg	在几个示例中使用的小鼠侏儒型基因

Q	纯合基因型 A_2A_2 的频率
q	对性状具减效作用的等位基因 A_2 的基因频率
R	选择反应——特指对个体选择
R_T	[12]总范围;即处于相反选择极限的两个群体之间平均值的差异
r	[8]重复力;即同一个体的重复度量之间的相关
r	亲缘系数;即亲缘个体之间育种值的相关
r	[10]配偶之间的表型相关
r	[19]两个性状之间的相关:r_A = 育种值的相关,r_E = 环境相关,r_P = 表型相关
S	有实际度量单位的选择差
SCA	特殊配合力
s	对一个特定基因型不利的选择系数(比例)
s	[13]下标,指同胞选择
T	各种含义的总数
t	以世代数表示的时间。作为下标指"在第 t 世代"
t	家系成员之间的表型相关(组内)
u	突变率(从 A_1 到 A_2)
u	[9,15]亲属协方差中显性方差的系数
V	由下标表明的离差值方差(有因组分):V_P = 表型方差,V_C = 基因型方差,V_A = 加性遗传方差,V_D = 显性方差,V_I = 互作方差(上位),V_{NA} = 非加性遗传方差,V_E = 环境方差
V_M	在一个世代中由突变产生的方差
V_S	[20]稳定选择强度的一种度量
v	突变率(从 A_2 到 A_1)
W 或 w	作为下标,表示家系内或组内
W	[20]自然选择下的适合度
X	下标,表示任一特定个体,例如,[5] F_X = 个体 X 的近交系数
X	两个相关性状之一
x	[11,18]正态离差;以标准差为单位表示的截断点距群体平均值的离差

Y	两个相关性状中的另一个
y	两个系群之间基因频率的差异
z	以标准差为单位表示的正态分布纵坐标的高度
α	基因替代的平均效应
α_1, α_2	分别表示等位基因 A_1 和 A_2 的平均效应
Δ	改变量，如 Δq = 基因频率的变化，ΔF = 近交率
$\sum$	对该符号后面的数量求和
σ	由下标表示的数量的标准差（σ^2 = 方差）

Mather 和 Jinks 使用的符号的等价意义

Mather 和 *Jinks*	本书
d	a
$[d]$	$\sum a$
D	当所有 $p = q = \frac{1}{2}$ 时，$\sum a^2 = 2V_A$（等式[8.7]）
D_R	在随机交配群体中 $2V_A$
E_w	V_{Ew}
E_b	V_{Ec}
E_1	V_{Ew}
E_2	$V_{Ec} + \frac{1}{n}V_{Ew}$
h	d
$[h]$	$\sum d$
H	当所有 $p = q = \frac{1}{2}$ 时，$\sum d^2 = 4V_D$（等式[8.7]）
H_R	在随机交配群体中 $4V_D$
u	p
v	q

习题解答

［方括号中的数字指的是在正文中编号的等式。］

1 （1.1）

（1）计数的总数是 6 129。对这样大的数字，频率需要 5 位小数位置来避免四舍五入的误差。基因型频率是：

MM：$P = 1\ 787/6\ 129 = 0.29156$

MN：$H = 3\ 039/6\ 129 = 0.49584$

NN：$Q = 1\ 303/6\ 129 = 0.21260$

检查这些频率加起来等于 1。

（2）将 P、H 和 Q 的数字代入［1.1］，给出基因频率为：

M：$p = [1\ 787 + (\frac{1}{2} \times 3\ 039)] / 6\ 129 = 0.53948$

N：$q = [1\ 303 + (\frac{1}{2} \times 3\ 039)] / 6\ 129 = 0.46052$

检查 $p + q = 1$

（3）根据［1.2］，期望的基因型频率是：

MM：$(0.53948)^2 = 0.29104$

MN：$2 \times 0.53948 \times 0.46052 = 0.49688$

NN：$(0.46052)^2 = 0.21208$

检查期望频率加起来等于 1。

（4）很紧密的一致。为了检验，我们必须将期望频率转换成可与观察数目比较的期望数目。将（3）中的频率乘以总数，给出期望数目为：

M	MN	N
1 783.8	3 045.4	1 299.8

检查数目加起来等于 6 129。根据 $\sum$［（观察值 − 期望值）^{2}/期望值］计算 χ^2，从中可得 $\chi^2 = 0.027$。这个很低的值证实了紧密的一致。这个 χ^2 有一个自由度，因为观察数目用来估计基因频率，并且期望数目必须与这个观察数及总数一致；换句话说，有 3 个数 2 个受约束，因此剩下 1 个自由度。

2（20.1）嗓深（Y）的观察选择差是继适合度（W）选择之后的一个相关选择差 S'。所做预测是 $R = h^2 S'$。但是，反应是一个相关反应，应当根据［19.8b］从 S' 中来预测：

$$CR_Y = \frac{r_A}{r_P} h_W h_Y S'$$

因此，通过 R 可以正确地预测 CR_Y，如果：

$$\frac{r_A}{r_P} h_W h_Y = h_Y^2$$

即，如果

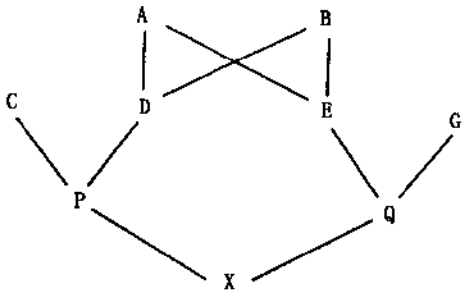
$$\frac{r_A}{r_P}=\frac{h_Y}{h_W}$$

3 （3.1）如果基因型取自一个随机交配群体，则基因型将处于哈迪—温伯格比例。像习题 1.1 中那样计算，这些是：

aa	*ab*	*bb*	总数
101.0	821.0	1 669.0	2 591

观察到的数目表明两种纯合体过量，杂合体相应的不足。这个差异是高度显著的。（$\chi_1^2=12.9$，$P<0.001$。）数据表明群体是基因频率不同的亚群体混合物。

4 （5.1）下面一单个嫡表兄妹的婚姻系谱将用来说明所有 3 种亲缘关系。



根据［5.1］或根据［5.2］，可由两种方法获得解答。将［5.1］应用于系谱：

（1）通径是 PDAEQ 和 PDBEQ，给出 $F_X=(\frac{1}{2})^5+(\frac{1}{2})^5=\frac{1}{16}$。

（2）现在 C 和 G 还是全同胞，因此还有 $n=5$ 的两条额外的通径，给出 $F_X=4\times(\frac{1}{2})^5=\frac{1}{8}$。

（3）假定 D 是叔父，他与他的侄女结婚。通径是 DAEQ 和 DBEQ；$F_X=2\times(\frac{1}{2})^4=\frac{1}{8}$。

为将［5.2］应用于（1）和（2），我们需要 4 种共祖率 CE、CG、DE、DG。对（1），除 DE 外，全为 0，因为 D 和 E 是全同胞，所以它们的共祖率是$\frac{1}{4}$。$F_X=\frac{1}{4}(\frac{1}{4})=\frac{1}{16}$。对于（2），CG 也是$\frac{1}{4}$；$F_X=\frac{1}{4}(\frac{1}{2})=\frac{1}{8}$。对于（3），所需要的共祖率是 AE，AG，BE，BG，它们当中 AE 和 BE 是$\frac{1}{4}$，是亲本和后代，其余的为 0；$F_X=\frac{1}{8}$。

5 （6.1）值的范围足够小以至于观察值不必集合成更宽的组。为了构建这张表，使叶片数目按次序成为一纵行，它们位于表中 X 下面。然后搜寻数据，根据叶片数目为每一植株作一记号。为了易于计数，每划 4 条斜线后再划一横线（相当于划“正”字），每 5 个为一组，以便快速计数。最后汇总计数，它们位于表中 n 下面。平均值和方差的计算列在表的底部。如果叶片数目被编码为偏离 12 的离差，像 x 下面展示的那样，可以简化计算。主要差异是 F_2 具有较大的可变性。

X	x	x^2	F_1			F_2		
			n	nx	nx^2	n	nx	nx^2
12	0	0	—	—	—	1	0	0
13	1	1	1	1	1	3	3	3
14	2	4	2	4	8	5	10	20
15	3	9	7	21	63	4	12	36
16	4	16	11	44	176	3	12	48
17	5	25	1	5	25	3	15	75
18	6	36	3	18	108	2	12	72
19	7	49	—	—	—	1	7	49
20	8	64	—	—	—	2	16	128
21	9	81	—	—	—	1	9	81
Σ	—	—	25	93	381	25	96	512
平均值				15.72 ± 0.24			15.84 ± 0.49	
方差				1.46			5.97	

平均值 $= 12 + \frac{\sum nx}{25}$　　方差 $= \frac{1}{24}\left[\sum (nx^2) - \frac{(\sum nx)^2}{25}\right]$

平均值的标准误 $= \sqrt{\text{方差}/25}$

6 （7.1）将频率乘以活性，对基因型求和，（如果采用频率的百分率）将总数除以100。

	频率 × 活性
AA	1 171.2
AB	7 438.2
BB	6 448.4
AC	515.2
BC	1 060.0
平均值 =	16 633.0/100 = 166.33

7 （8.1）第一个假设是杂交的品种在所有位点上都是纯合的。烟草正常是自花传粉的，因此这很可能是真实的。于是从起源上看，F_1 方差完全是环境的。F_2 方差在起源上既有环境的又有遗传的。第二个假设是 F_2 中的环境方差与 F_1 中的相同。借助这个假设，遗传方差可通过减法而获得。

F_2 方差 $\simeq V_G + V_E = 5.97$

F_1 方差 $= V_E = 1.46$

$V_G = 4.51$

遗传决定的程度 $= \frac{V_G}{V_G + V_E} = \frac{4.51}{5.97} = 0.76$ 或 76%。

8 （9.1）孩子们是半同胞（见习题5.2），因此根据表9-3，$r = \frac{1}{4}$。

9 （11.1）利用附录表A，从选择比例 p 中可得到选择强度 i。然后将 σ_p 作为方差的平方根，应用等式［11.3］。解答如下。

（1）（a）$R = 1.271 \times 0.37 \times 3.271 = 1.54$g。

（b）$R = 0.798 \times 0.37 \times 3.271 = 0.97$g。

(c) 当 p 大于 50% 时，取 $1-p$ 的 i 值，并乘以 $(1-p)/p$。$i=1.271\times0.25/0.75=0.424$；$R=0.51$g。

(2) $R=1.755\times0.18\times1.304=0.41$ 天。

(3) 对雌性的选择强度 = 1.159。对雄性没有进行选择，因此 $i=\frac{1}{2}\times1.159=0.5795$；$R=0.5795\times0.22\times2.074=0.26$ 头仔畜/窝。当对生殖力进行选择时，后代已经出生，而此时已度量生殖力的个体就可以被选择了。考虑这个过程的一个备择方法是，将两个性别的后代看作是在其母亲生殖力的基础上选择的。后代对母亲的回归是 $\frac{1}{2}h^2$，因此在这个例子中反应是 $R=1.159\times(\frac{1}{2}\times0.22)\times2.074=0.26$ 头仔畜/窝。

10 (17.1) 方差可借助 [17.2] 转换成对数，结果如下。

	小 (S)	对照 (C)	大 (L)
C	0.143	0.111	0.128
$1+C^2$	1.0204	1.0123	1.0164
$\sigma^2_{(对数)}\times100$	0.3809	0.2306	0.3068
$\sigma_{(对数)}$	0.062	0.048	0.055
对数的平均值	1.074	1.362	1.597
对数的反应	0.288	0.235	

当反应转换成对数时，如果反应相等，则算术平均值的比率将相等：

L/C = 39.85/23.16 = 1.72

C/S = 23.16/11.97 = 1.93

不对称程度被大大减少。可选择的是，对数转换数据的平均值可从 [17.1] 中 $(\overline{\log x})$ 的公式中计算出来。上面给出了这些值。用对数单位表示的反应几乎是相等的。这从下面的事实中可预测出来：反应与标准差成比例 ([11.3])，并且对数单位的标准差几乎是相等的。

11 (1.2) 对非品尝基因非品尝员是纯合体。具有哈迪—温伯格频率，纯合体的频率是基因频率的平方。因此 $q=\sqrt{0.3}=0.55$。

12 (2.1) 首先得到不利于白花植株的选择系数 s。相对于蓝色的适合度是 143/229 = 0.62，$s=1-0.62=0.38$。假定随机授粉，白花植株是纯合体，它们的频率是 q^2。根据 [2.12]，突变率是：

$$u=sq^2=0.38\times7.4\times10^{-4}=2.8\times10^{-4}$$

注意适合度 0.62 是雌性植株的适合度。计算过程假定雄性的适合度同样地降低。

13 (3.2) 单独计算每一个族的基因频率和哈迪—温伯格期望值，并与观察数目比较。χ^2 检验观察和期望数目之间的一致性。

	(a 的) p	aa	ab	bb	χ^2_1
北冰洋	0.1214	18.0	260.0	941.0	1.8
海岸	0.2649	96.3	534.3	741.4	2.2

在两个族中基因频率不同，因此在混合样本中预料有过量的纯合体。每一个族，对它自己

的基因频率来说，具有与哈迪—温伯格期望值良好一致的基因型频率。因此细分成这两个族，每个族都是族内随机交配，足以解释总样本中的基因型频率。

14 （5.2）双生子在遗传上等价于一单个个体。因此这两个婚姻等价于一个个体具有两个配偶，因此孩子为半同胞相关，他们的共祖率是 1/8。（见［5.7］。）如果根据［5.1］从系谱中来求解，则这对双生子在系谱中必须视为一单个个体。

15 （6.2）数据必须组群成几类。以 1g 为间隔组合成 15 类，这是令人满意的。类别 X 根据重量的整数部分制成表，即忽略小数部分。类别间隔是 4.0～4.9，5.0～5.9，等等，类别中点是 4.45，5.45，等等。因此平均值是 $0.45+\overline{X}$。不同的组群将产生稍微不同的平均值和方差的估计值。

X	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
n	1	2	5	1	0	0	1	6	4	8	10	7	2	2	1

	全部	侏儒	正常
平均值	12.61	6.12	14.04
方差	12.34	0.75	3.40

分布是双峰的，没有重叠。9 个很小的小鼠在某一方面明显不同于其余的。事实上，它们对称为侏儒的一种矮化基因是纯合的，第 7 章和第 8 章中这种基因被用作例子。侏儒基因的存在大大增加了方差。

16 （7.2）运用［7.2］能够做到这一点，但通过计算哈迪—温伯格频率和通过对活性×频率求和得到平均值，可能更简单。

	活性	频率		
		(1)	(2)	(3)
AA	122	0.04	0.25	0.64
AB	154	0.32	0.50	0.32
BB	188	0.64	0.25	0.04
平均值		174.48	154.50	134.88

17 （8.2）这仅需要将 α、d 和 q 值代入［8.3a］和［8.4］中。α 和 d 的值在解第 7 章习题中已获得。

		α	d	q	$2pq$	$V_A=2pq\ \alpha^2$	$V_D=(2pqd)^2$
(1)	(1)	33.6	-1	0.2	0.32	361.27	0.10
	(2)	33.0	-1	0.5	0.50	544.50	0.25
	(3)	32.4	-1	0.8	0.32	335.92	0.10
(2)		0	50	0.4	0.48	0	576.00
(3)	基因 b:	2.5	2.5	0.5	0.50	3.125	1.56
	基因 c^s:	11.4	28.5	0.2	0.32	41.59	83.17
	两个基因:					44.71	84.74

18 （9.2）习题 8.2（3）的解答给出了 $V_A=44.71$ 和 $V_D=84.74$，我们必须把 $V_E=\frac{1}{3}(V_A+V_D)=43.15$ 加到它们上。这 3 个组分相加给出 $V_P=172.60$。根据表 9-3 中的协方差，于是我们获得：

(1) 后代对中亲的回归 $= V_A/V_P$ $= 0.259$

(2) 后代对单亲的回归 $= \frac{1}{2} V_A/V_P$ $= 0.130$

(3) 全同胞相关 $= (\frac{1}{2} V_A + \frac{1}{4} V_D) / V_P$ $= 0.252$

(4) 半同胞相关 $= \frac{1}{4} V_A/V_P$ $= 0.065$

(5) 双重嫡表兄妹相关 $= (\frac{1}{4} V_A + \frac{1}{16} V_D) / V_P$ $= 0.095$

19 (11.2) 由于在第 19 章和第 20 章要解释的原因，不作出一个重要假设，这个问题就不能回答，即选择存活的原因是喙深本身而不是与它相关的某个其他性状，例如翅长。这个假设在所描述的环境中似乎是有理由的。于是，借助这个假设，选择差是 $S = 9.96 - 9.42 = 0.54$mm。为了预测反应，我们需要知道遗传力，遗传力将取自习题 10.8，其大小为 0.82。根据 [11.2]，预测反应是 $R = 0.82 \times 0.54 = 0.44$mm，增加了 5%。对生存者后代预测的平均值是 $9.42 + 0.44 = 9.86$mm。所作的假设是习题 20.1 的主题。

20 (17.2) 依杂种和中亲值之间的差异计算杂种优势。首先对给定的算术值做这一点。然后将所有给定的算术值转换成对数并从这些对数中计算杂种优势。在两个尺度上的杂种优势展示在下面。我们不能使用 [17.1] 来评估对数值平均值，因为变异系数没有给出，因此通过对给定算术平均值取对数，尺度转换一定做得不够准确。这里采用底数为 10 的对数。使用自然对数 $\log_e$，可同样很好。两者之间的关系是 $\log_{10} x = 0.4343 \log_e x$。

	算术				对数		
	L	C	S		L	C	S
L	1.94	1.135	0.025	L	0.028	0.025	0.023
C	—	0.92	0.53	C	·	0.019	0.019
S	—	—	-0.07	S	—	—	-0.002

根据杂交系群的大小，在算术尺度上杂种优势变化很大；不容易看出大小组之间的杂种是否不同于大小组之内的杂种。在对数尺度上，杂种优势在大小组之间和大小组之内的所有杂种中几乎都相同，除 S×S 杂交以外，它在两种尺度上都是不规则的。两种尺度上杂种优势绝对大小的差异没有意义。

21 (1.3) 对一个近似的答案，如果频率被表示为分数，则求解较容易。首先从 $q^2 = 1/20\,000$ 中得到基因频率；$q = 1/141$。携带者是杂合体，具有频率：

$$2q(1-q) = 2 \times \frac{1}{141} \times \frac{140}{141}$$

$$= \frac{2}{141} \text{（近似）}$$

$$= \frac{1}{70} \text{（近似）}$$

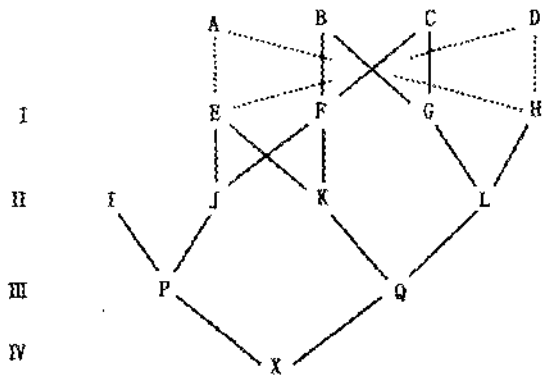
22 (2.2) 白花纯合体极端稀少，可以忽略。于是，白花植株的频率是杂合体的频率 H。根据 [2.15]，从蓝到白的突变率是 $u = sH/2 = \frac{1}{2}(0.38 \times 7.4 \times 10^{-4}) = 1.4 \times 10^{-4}$。

23 (3.3) $N=40$。根据 [3.7]，$\Delta F=1/80=0.0125$

根据 [3.12]，$F\ (t=5)\ =1-\ (0.9875)^5=0.061$

$F\ (t=10)\ =1-\ (0.9875)^{10}=0.118$

24 (5.3) 首先重画系谱以便更清楚地显示传递通径，并标注上世代 I 的亲本，这样：



有 6 条通径，2 条通过共同祖先 E 和 F，$n=5$，4 条通过 A、B、C 和 D，全都具有 $n=7$。6 个共同祖先没有一个是近交种。根据 [5.1]

$$\begin{aligned}F_X&=2\ (\tfrac{1}{2})^5+4\ (\tfrac{1}{2})^7\\&=\ (\tfrac{1}{2})^4+\ (\tfrac{1}{2})^5\\&=\frac{1}{16}+\frac{1}{32}\\&=\frac{3}{32}\\&=0.09375\end{aligned}$$

亲本 Q 是双重嫡亲兄妹结婚所生的孩子，$F=\frac{1}{8}$，这一点可从习题 5.1 中看出。注意亲本的近交不影响孩子的近交。

25 (6.3) 有 5 个基因型类别，根据隐性等位基因纯合的位点数，其度量值是 -4、-3、-2、-1、0。这些类别的频率由 $(a+b)^4$ 的二项式展开的各项给出，其中 a 是在任一特定位点上纯合的概率。这里 $a=(0.3)^2=0.09$，$b=0.91$。二项式展开的各项是：

n								
1					1		1	
2				1		2		1
3			1		3		3		1
4		1		4		6		4		1
5	1		5		10		10		5		1
6	1	6		15		20		15		6		1

$$a^4 \qquad 4a^3b \qquad 6a^2b^2 \qquad 4ab^3 \qquad b^4$$

系数 1，4，6，4，1 最容易从上面给出的一直到 $n=6$ 的具有 $n=4$ 的杨辉三角中获得。达到 3 位小数所要求的频率是：

度量类别	-4	-3	-2	-1	0	总数
频率	0.000	0.003	0.040	0.271	0.686	1.000

极端的不对称性是隐性纯合体低频率的后果。

26 (7.3) 这可以通过 3 种方法来做到。(i) 根据 [7.2] 计算几个不同基因频率的 M 并以 q 为背景用图表示 M。当然，这只能给出一个近似答案。

(ii) 处理数量值就好像它是适合度，利用 [2.19] 来找出平衡基因频率。为此，我们需要每一个纯合体相对于杂合体的适合度。$(1-s_1)\ =110/150=0.733$，$(1-s_2)\ =90/150=0.600$；$s_1=0.267$，$s_2=0.400$。于是，如果 $\hat{q}$ 是 A_2 的平衡频率，根据 [2.19]，$\hat{q}=$

$0.267/0.667 = 0.4$。

(iii) 据 [7.2] 对 q 求微分并使其等于 0。(在使用 [7.2] 过程中，必须小心确保 p 是给予较高值的等位基因的频率，在这个例子中是 A_1。) 将 $p = 1 - q$ 代入 [7.2]，重新整理给出：

$$M = a + 2(d - a)q - 2dq^2$$

$$dM/dq = 2(d - a) - 4dq$$

$$\hat{q} = \frac{d - a}{2d}$$

纯合体中值是 100，$a = 10$，$d = 50$。这样 $\hat{q} = 40/100 = 0.4$。

将这些 a、d 和 $\hat{q}$ 的值代入 [7.2]，给出 $M = 26$，作为偏离纯合体中值的离差。因此群体平均值是 $100 + 26 = 126$。

27 (8.3) 比率 V_D/V_G 是 [8.4] 除以 [8.8]。让 c 代表显性度，因此 $d = ca$。然后 $2pqa^2$ 相消，比率简化成：

$$\frac{V_D}{V_G} = \frac{2pqc^2}{2pqc^2 + [1 + c(q - p)]^2}$$

其中 q 是隐性等位基因的频率。对不同基因频率求解之前，代入 c 值然后进一步简化要更容易一些。下面给出了简化表达式，具有 4 种基因频率的 V_D/V_G 的值。为了绘制曲线，将必须计算更多的值。

	$d = \frac{1}{2}a$	$d = a$	$d = 2a$
q	$\frac{q - q^2}{q^2 + 3q + \frac{1}{2}}$	$\frac{1 - q}{1 + q}$	$\frac{q - q^2}{q^2 + \frac{1}{8}}$
0.2	0.140	0.667	0.970
0.4	0.129	0.429	0.842
0.6	0.090	0.250	0.495
0.8	0.045	0.111	0.209
最大值	0.143	→1	1.000
在 q =	0.250	→0	0.250

注意由低频率的一个完全隐性等位基因造成的大部分方差是显性方差；但当显性不完全时，显性方差的比例不是很大。

28 (9.3) 父子和母子相关显然是一致的。为了与中亲—孩子相关进行比较，我们必须将相关转换成回归。在做这个的过程中，我们必须作两个假设：亲本的方差和孩子的方差相同，孩子是单个孩子，不是几个孩子的平均值。借助这些假设，孩子对单个亲本的回归等于相关，都是$\frac{1}{2}V_A/V_P$ 的估计值，V_A/V_P 的估计值是 1.00 和 0.98。中亲—孩子相关估计了 $(\sqrt{\frac{1}{2}})$ V_A/V_P，给出 $V_A/V_P = 0.69/\sqrt{\frac{1}{2}} = 0.98$。因此这 3 种相关完全是一致的。

29 (11.3) 选择强度可从附录表 B 中获得；它在两个性别中不同。对雄性，$n = 4$，$N = 60$，给出 $i_m = 1.882$。对雌性，$n = 8$，$N = 60$，$i_f = 1.582$。平均值 (见 [11.6b]) 是 i

= 1.732。根据［11.3］，每世代的预测反应是 $R = 1.732 \times 0.81 \times 111 = 155.7$g。

如果反应以相同的速度继续下去，那么 5 代之后平均值将是 738 + （5 × 155.7） = 1 517g。可是，对这样高的遗传力，由选择引起的方差的减少是不能忽略不计的。通过取平均中选比例，即 10% 可以作出很粗略的预测。于是表 11-6 中的内插法给出 $R_2 = 0.78 \times R_1 = 121$g。将随后的小的反应减少忽略，5 个世代之后的预测平均值是 738 + 155.7 + （4 × 121） = 1 378g。

30 （17.3）因为效应随平均值而增加，首先我们可以尝试一种对数转换；或更简单地，考察比率 H/L 来看看它是否稳定不变。在转换成对数之后 *X* 染色体之间的差异给在下表的左边。它对算术差异是一个相当的改进，但它仍然随平均值而增加。因此我们需要一种更强烈的转换。在转换成对数之前，通过减去某一常数可以达到这个目的。人们可能猜想应当减去 4 根较大的刚毛，因为它们几乎是不变的。于是转换将是 log（$x-4$），其中 x 是计数的刚毛数。这种转换效果很好，正如下面所示，提供了独立于常染色体水平的 *X* 染色体的效应。

	对数差异	log（$x-4$）		
		高	低	差异
A	0.088	0.740	0.574	0.166
B	0.090	0.970	0.835	0.135
C	0.143	1.272	1.092	0.180
D	0.147	1.490	1.319	0.171
E	0.164	1.641	1.459	0.182

可以更合理地获得所要求的转换，结果如下。在算术纸上绘制高对低算术值作图。这些点几乎位于一条直线上。我们在轴上要寻找一个尺度，使得直线通过原点，即在两个轴上 0 点处。这将给出 $y = bx$ 形式的一种图形，这样比率 y/x 是稳定的。如果绘制的直线向下伸展，将发现它通过或接近 H = 4，L = 4 这一点。这意味着（H − 4）/（L − 4）是稳定的，所要求的转换就是 log（$x-4$）。

31 （1.4） $\dfrac{\text{杂合体}}{\text{正常个体}} = \dfrac{2pq}{1-q^2} = \dfrac{2q(1-q)}{(1+q)(1-q)} = \dfrac{2q}{1+q}$

代入

$$\frac{2q}{1+q} = \frac{1}{3}$$

给出解，$q = \dfrac{1}{5} = 0.2$。

（注意：记住当两个等位基因的基因频率记为 p 和 q 时，$p + q = 1$，因此 $p = 1 - q$。在做基因频率的代数时，代入 $p = 1 - q$，或 $q = 1 - p$ 作为第一步通常是最好的。）

32 （2.3）根据［2.4］，a 的基因频率将是：

$$10^{-4}/(10^{-4} + 10^{-5}) = 1/(1 + 10^{-1}) = 10/11 = 0.90909$$

于是哈迪—温伯格频率是 $AA = 0.0083$，$Aa = 0.1653$，$aa = 0.8264$。

33 （3.4）期望平均值是基础群体的基因频率，即 0.3。如果学生数很大，则将发现这个。经过电泳的样本可视为第六世代的亲本，因此就有 6 个世代的随机漂变。首先像习题 3.3 那样得到 F（$t = 6$），但 $N = 20$。这给出 $F = 0.141$。然后，根据［3.14］，基因频率的方差将是 $\sigma_q^2 = 0.3 \times 0.7 \times 0.141 = 0.030$。学生估计值的标准差将是 $\sqrt{0.030} = 0.17$。

34 （5.4）首先根据［5.15］得到近交系数为：

$$F = \frac{1 - 0.01}{1 + 0.01} = 0.9802$$

然后，根据［3.15］，杂合体相对于哈迪—温伯格频率的频率是：

$$H_t/H_0 = 1 - F = 0.0198$$

哈迪—温伯格频率是：

$$H_0 = 2 \times 0.2 \times 0.8 = 0.32$$

因此，在群体中所要求的杂合体频率是：

$$H_t = 0.0198 \times 0.32 = 0.63\%$$

35 （6.4）首先必须对两种位点分别计算出二项式频率。在每一种情形中都有 3 种基因型类型，具有二项式频率 $a^2 \quad 2ab \quad b^2$，其中 a 和 b 如下。

$q = 0.3$ 的位点：　　$a = 0.09$，　　$b = 0.91$

$q = 0.7$ 的位点：　　$a = 0.49$，　　$b = 0.51$

两种位点的二项式频率和基因型类型列在下表的边际（顶部和左边）。

	类型	频率	$q=0.3$ 的位点		
			−2	−1	0
			0.008	0.164	0.828
$q=0.7$ 的位点	−2	0.240	−4	−3	−2
			0.002	0.039	0.199
	−1	0.500	−3	−2	−1
			0.004	0.082	0.414
	0	0.260	−2	−1	0
			0.002	0.043	0.215

两种类型的位点放在一起位于表的中央部分。基因型类型通过将两个边际类型加起来而得到，频率通过将两个边际频率相乘而得到。有 5 个基因型类型。将代表同一类型的小区中的频率加在一起给出如下的频率分布。

度量类型	−4	−3	−2	−1	0	总数
频率	0.002	0.043	0.283	0.457	0.215	1.000

36 （7.4）像图 7-1 中定义的那样，基因型值是 $a_b = \frac{1}{2}(95 - 90) = 2.5$；$a_c = \frac{1}{2}(95 - 38) = 28.5$。（下标 c 表示 c^e 基因。）在两种情形中，$d = a$、于是［7.2］简化成 $M = a(1 - 2q^2)$，其中 q 是隐性和减效等位基因的频率。这给出：

$$M_b = 2.5\,[1 - 2\,(0.5)^2] = 1.25 \quad \text{和} \quad M_c = 28.5\,[1 - 2\,(0.2)^2] = 26.22$$

像［7.3］那样，将两个基因放在一起，$M = 1.25 + 26.22 = 27.47$。这是偏离两个双纯合体中点的离差。已经表述，双显性纯合体的值是 95。假定基因效应是加性的，那么双隐性纯合体的值是：

$$95 - 2a_b - 2a_c = 95 - 5 - 57 = 33$$

因此中间纯合体值是$\frac{1}{2}(95 + 33) = 64$。这使得平均颗粒数为 $64 + 27.47 = 91.47$。

得到平均值的另一种可选择的和更快的方法是来自 4 种基因型的值和频率，正如上表

基因型		B-	bb
	频率	0.75	0.25
C-	0.96	95 0.72	90 0.24
c^ec^e	0.04	38 0.03	34 0.01

所示。列入的值，即颗粒数是例 7.3 中给出的数值。

平均值 = (95 × 0.72) + (90 × 0.24) + (38 × 0.03) + (34 × 0.01) = 91.48。

37 (8.4) 对每一个基因，我们必须分别通过 [8.3a] 获得 V_A，通过 [8.4] 获得 V_D，然后将它们加在一起。做这个所需的基因型值 a 和 d 必须从习题 7.9 解答中的表 (ii) 边上的平均值中取得。

基因 b: $a = d = \frac{1}{2}(0.228 + 0.342) = 0.285$; $q = \sqrt{0.4} = 0.6325$

基因 c^e: $a = d = \frac{1}{2}(0.060 + 0.240) = 0.150$; $q = \sqrt{0.2} = 0.4472$

基因 b: $\alpha = 0.285\ (1 + 0.6325 - 0.3675) = 0.3605$

基因 c^e: $\alpha = 0.150\ (1 + 0.4472 - 0.5528) = 0.1342$

	$V_A = 2pq\alpha^2$	$V_D = (2pqd)^2$
基因 b	0.0604	0.0176
基因 c^e	0.0089	0.0055
两个基因	0.0693	0.0231

互作方差 V_I 直接从习题 7.9 解答的表 (iv) 的互作离差中计算出来。表中的值是偏离群体平均值的离差，因此它们的方差简单地就是它们平方的平均值。因此为了得到 V_I，将每一个互作离差的平方乘以基因型频率，再相加：

$V_I = 0.48\ (0.04)^2 + 0.32\ (-0.06)^2 + 0.12\ (-0.16)^2 + 0.08\ (0.24)^2 = 0.0096$

$V_A = 0.0693 = 67.94\%$

$V_D = 0.0231 = 22.65\%$

$V_I = 0.0096 = 9.41\%$

$V_G = 0.1020 = 100.00\%$

为了检验，可直接从习题 7.9 的表 (ii) 的值解出 V_G。表 (iii) 中加性期望值方差给出 $V_A + V_D$。

38 (9.4) 对标准化的窝仔畜，全同胞相关比母女回归稍高些。这是由显性或共同环境或两者共同作用引起的。由于下面的原因，被养育母畜的窝产仔数影响她自己随后的窝产仔数。在较大的窝中养育的母畜不得不与较大数量的竞争者合用它们出生前和断奶前的营养，因此它们在断奶时较小，发育到成年时也较小。因为较小，它们排出较少的卵，具有较小的窝仔畜。作为窝同伴被养育的同胞分享它们被养育的窝仔畜环境。因此当窝仔畜没有进行标准化时，窝产仔数的变异造成了增加全同胞相关的环境协方差，这一点可从数据中看出。被养育的一头母畜所处的窝大小是母亲的窝产仔数。因此，当窝仔畜没有被标准化时，具有较大窝仔畜的母亲倾向于产生具有较小窝仔畜的女儿。因此，女儿和母亲的窝产仔数具有一种负的环境协方差。当窝仔畜没有进行标准化时，这抵消了正的遗传协方差，结果母女回归几乎为 0。注意这种母体效应如何引起窝同伴的正环境协方差，又是如何引起后代和亲本的负环境协方差。

39 （11.4）首先必须计算出每一个世代的选择差：$S = P - M$。这是选择，对它的反应在下一世代才能看出来。因此每一个选择差依靠后裔世代进入表中。世代 0 没有应用于它的选择。表中位于反应 R 下的数字简单地就是世代平均值。下一步，选择差连续地加起来以给出在 $\sum S$ 之下所显示的累积选择差。对于差异，每一个世代中的 R 和 $\sum S$ 值是通过从高系的值中减去低系的值而获得的。通过 R 对 $\sum S$ 的回归来估计现实遗传力，包括世代 0 的值在内。另外可选择的是，虽然不够可靠，用总反应与总选择的比率来估计它。对于高系，这是 $(2.44 - 2.16) / 0.45 = 0.62$。

世代	高			低			差异	
	R	S	$\sum S$	R	S	$\sum S$	R	$\sum S$
0	2.16	0	0	2.16	0	0	0	0
1	2.26	0.16	0.16	2.06	-0.14	-0.14	0.20	0.30
2	2.26	0.08	0.24	2.03	-0.06	-0 20	0.23	0.44
3	2.33	0.11	0.35	2.02	-0.06	-0.26	0.31	0.61
4	2.45	0.08	0.43	2.05	-0.06	-0 32	0.40	0.75
5	2.44	0.02	0.45	2.01	-0.04	-0.36	0.43	0.81
R 对 $\sum S$ 的回归			0.631			0.362		0.512
总反应 总选择			0.62			0.42		0.53

40 （19.1）遗传力（见表 10-7）

增重：$h_G^2 = 4 \times \dfrac{1\,602}{12\,321} = 0.52$

饲料消耗量：$h_F^2 = 4 \times \dfrac{6\,150}{61\,504} = 0.40$

表型相关：$r_P = \dfrac{22\,848}{\sqrt{12\,321 \times 61\,504}} = 0.83$

遗传相关：$r_A = \dfrac{2\,229}{\sqrt{1\,602 \times 6\,150}} = 0.71$

环境相关：$r_E = \dfrac{22\,848 - (4 \times 2\,229)}{\sqrt{(12\,321 - 4 \times 1\,602)(61\,504 - 4 \times 6\,150)}} = 0.94$

可供选择的是，也可以根据［19.1］来计算环境相关，为此需要下面的数据

$h_G = 0.7211$　　$e_G = 0.6928$

$h_F = 0.6325$　　$e_F = 0.7746$

$h_G h_F = 0.4561$　　$e_G e_F = 0.5366$

代到［19.1］中给出：

$0.83 = (0.71 \times 0.4561) + (r_E \times 0.5366)$

$r_E = 0.94$

41 （1.5）基因可以计数，它们的频率可通过［1.1］的扩展来确定。

A：$0.096 + \dfrac{1}{2}(0.483 + 0.028) = 0.3515$

B：$0.343 + \dfrac{1}{2}(0.483 + 0.050) = 0.6095$

$$C: \qquad 0 \qquad + \frac{1}{2}(0.028 + 0.050) = 0.0390$$

总数 1.0000

CC的哈迪—温伯格期望值是 $(0.0390)^2 = 0.0015$。在178个人的样本中期望数目是0.27。没有发现一个，因为期望值大大低于1。

42 (2.4) 没有。根据 [2.4]，当两种速率增加同一比例时，平衡是相同的。

43 (3.5) 习题3.4给出 $\sigma_q^2 = 0.03$。起始基因频率是0.3和0.7。于是根据 [3.5]，基因型频率将是：

A_1A_1: $0.09 + 0.03 = 0.12$

A_1A_2: $0.42 - 0.06 = 0.36$

A_2A_2: $0.49 + 0.03 = 0.52$

44 (5.5) 解答来自重新整理的 [5.15]，用 F 解出 C。根据 [3.15] 可得到 F，但为此我们需要杂合体的哈迪—温伯格频率 H_0。根据 [1.1]，基因频率是0.6和0.4。因此 $H_0 = 0.48$。于是，由 [3.15]，

$1 - F = 0.12/0.48 = 0.25$;

$F = 0.75$。

重新整理 [5.15] 变成：

$$C = \frac{1-F}{1+F} = \frac{0.25}{1.75} = 0.143$$

证明的异花授粉频率是14%。

45 (6.5) 在这种情形中，为了从二项式展开中获得频率，我们必须对基因而不是对基因型做研究。因为无显性，因此度量可由增效等位基因的数目所确定，每一个增加1个单位的度量。对3个位点，基因型可有从0到6个增效等位基因，分成7组。频率由 $(p+q)^6$ 的展开式给出，其中 p 和 q 是基因频率0.4和0.6。

度量组	0	1	2	3	4	5	6	
频率	p^6	$6p^5q$	$15p^4q^2$	$20p^3q^3$	$15p^2q^4$	$6pq^5$	q^6	总数
频率%	0.4	3.7	13.8	27.6	31.1	18.7	4.7	100.0

46 (7.5) 首先从 [7.5] 中得到基因替代的平均效应 α 是最好的。为此我们必须像图7-1中定义的那样评估 a 和 d。从习题7.1给出的数据中获取基因型AA、AB和BB酶活性的值，a 和 d 可计算为：

$$a = \frac{1}{2}(188 - 122) = 33$$

$$d = 154 - \frac{1}{2}(188 + 122) = -1$$

实际上可能无显性，即 $d = 0$，但用 $d = -1$ 来求解答案，好像它是真实的。在习题7.2中指定的基因频率 q 是等位基因A的频率，A是给予较低值的等位基因，这是在 [7.5] 中所要求的。

(1) $q = 0.2$。代入到 [7.5] 给出：

$$\alpha = 33 + [-1(0.2 - 0.8)] = 33 + 0.6 = 33.6$$

根据［7.6］可分别给出等位基因的平均效应。这里 α_1 指的是频率为 p 的等位基因，它是B。因此，根据［7.6］

$\alpha_B = 0.2 \times 33.6 = 6.72$

$\alpha_A = -0.8 \times 33.6 = -26.88$

检查 $\alpha_B - \alpha_A = \alpha$

（2）$q = 0.5$：　$\alpha = 33 + 0 = 33$；　$\alpha_B = 16.5$；　$\alpha_A = -16.5$。

（3）$q = 0.8$：　$\alpha = 33 - 0.6 = 32.4$；　$\alpha_B = 25.92$；　$\alpha_A = -6.48$。

（注意：对严格地无显性（$d = 0$），α 等于 a，α 对所有基因频率都相同，但 α_A 和 α_B 依然取决于基因频率。）

47　（8.5）对（1），我们需要第一和第二窝之间的相关 r，对（2），我们需要第二窝对第一窝的回归 b 和两个平均值。让 X 代表第一窝，Y 代表第二窝。

$\sum X = 104$　　$\sum Y = 103$

$\sum X^2 = 1\ 106$　　$\sum Y^2 = 1\ 101$　　$\sum XY = 1\ 089$

$\overline{X} = 10.4$　　$\overline{Y} = 10.3$

校正的平方和以及乘积和为：

$\sum x^2 = 24.4$　　$\sum y^2 = 40.1$　　$\sum xy = 17.8$

重复力被估计为乘积瞬时相关（product moment correlation），它是：

$r = 17.8/\sqrt{24.4 \times 40.1} = 0.57$

来自这个小样本的估计值比在较大样本中正常发现的要高。重复力应当估计为组内相关，可以算出为 0.59，但这不是严格有效的，当 X 和 Y 的方差不同时，它们确实不同：

$\sigma_X^2 = 2.7$，$\sigma_Y^2 = 4.5$。

（2）第二对第一窝产仔数的回归是：

$b_{YX} = 17.8/24.4 = 0.73$

第二窝的期望大小：　$(a)10.3 + 0.73(14 - 10.4) = 12.9$

$(b)10.3 + 0.73(5 - 10.4) = 6.4$

这些是（a）第一窝的全部小鼠是 14 只仔鼠，（b）第一窝的全部小鼠是 5 只仔鼠时对第二窝预测的平均大小。

48　（10.1）按照［10.5］，相关或回归必须乘以 $1/r$，r 的值给在表 9-3 中。在（4）和（7）的情形下，相似于 $1/r$ 的因子在正文中有解释并表示在下面的括号中。

	$1/r$	h^2
(1)	2	0.42
(2)	2	0.54
(3)	2	0.68
(4)	(1)	0.32
(5)	4	0.08
(6)	4	0.12
(7)	(2)	0.18

49　（11.5）未加权的选择差是亲本平均值和它们世代平均值之间的差异；计算展示

在下表中。注意选择是为了降低大小，因此具有负号的选择差是期望得到的。为了获得加权选择差，计算亲本的加权平均值，用后代数目进行加权。如果 P 是亲本值，n 是相应的后代数，那么加权平均值是 $\sum nP/\sum n$。对雌性，这就是：

$$[(7.6\times 1)+(12.4\times 9)+\cdots]/69=917.6/69=13.30$$

其余部分的计算展示在表中。

	未加权		加 权	
	雌性	雄性	雌性	雄性
亲本平均值（a）	12.59	13.01	13.30	13.75
亲本世代平均值（b）	13.14	14.80	13.14	14.80
选择差（a－b）	－0.55	－1.79	＋0.16	－1.05
性别平均值	－1.17		－0.445	

对雌性生殖力的自然选择与对小的身体大小的人工选择相反，在某种程度上，有效选择差是在错误的方向上。雄性中类似的但小得多的效应可能是偶然的。

50 （19.2）在一致群体中的方差和协方差只是环境的。从变化的群体值中减去这些值可以估计基因型的方差和协方差（见例 8.1）。用这种方法估计的遗传相关 r_G 是基因型值的相关。习题 19.1 中的遗传相关 r_A 是从公畜组分中估计出来的，是育种值的相关。

群体	变异原因	相 关
变化的	$G+E$	$0.87/\sqrt{(0.366\times 43.4)}=0.22=r_P$
一致的	E	$0.27/\sqrt{(0.186\times 16.6)}=0.15=r_E$
变化的－一致的	G	$0.60/\sqrt{(0.180\times 26.8)}=0.27=r_G$

51 （1.6）雄性的基因频率是患病雄性的频率，给定为 0.07。在哈迪—温伯格平衡下，雌性的基因频率与雄性的相同。（1）杂合妇女的频率是 $2q(1-q)=2\times 0.07\times 0.93=0.13$，或 13%。（2）色盲（即纯合）妇女的频率是 $q^2=0.0049$，或约 1/200。（3）结婚的频率是男人和女人的频率乘积，$0.07\times 0.0049=0.000343$，或大约 1/3 000。

52 （2.5）选择系数，起始为 $s=1$，因治疗而减半了。平衡时，现在和未来，根据 [2.13]，纯合体频率是 $q^2=u/s$。当达到新的平衡时，减半 s 将加倍纯合体频率。基因频率的增加来自突变，因此将花费很长时间才能达到新的平衡。

53 （3.6）观察到的杂合体频率是 $H_t=0.1343$。根据 [1.1]，作为一个整体在群体中基因频率是 0.1413 和 0.8587。在具有这样基因频率的一单个随机交配群体中杂合体的频率是 $H_0=0.2427$。根据 [3.15]，随机交配指数是：

$$P=\frac{0.1343}{0.2427}=0.553$$

近交系数是：

$$F=0.447$$

54 （5.6）根据 [4.1]，在第一期，$\Delta F=1/32$。根据 [3.11]，10 个世代之后，$1-F=P=(31/32)^{10}=0.728$。在第二期，$\Delta F=1/64$。在期末，将第二期的开始作为基础，$P=(63/64)^{10}=0.854$。将第一期的开始作为基础，根据 [5.17]，

$$P=0.854\times 0.728=0.622$$

$F = 0.378$

等价于全同胞交配两个世代（表 5-2）。

55 （6.6）当 $a = b = \frac{1}{2}$时，在 $(a+b)^n$ 中将导致一个对称分布。（1）对隐性，$a = q^2$，其中 q 是隐性等位基因的频率；$q = \sqrt{\frac{1}{2}} = 0.707$。（2）无显性时，$a = q$；$q = 0.5$。

56 （7.6）从表 7-9 中获得育种值，其中 A_1A_1 对应于 BB。在习题 7.5 中找出所需要的 α 值。

这些值是偏离群体平均值的离差。检查每一个群体中平均育种值为 0：将育种值乘以该群体中基因型频率（哈迪—温伯格），然后对基因型求和。

	(1) $q=0.2$ $\alpha=33.6$	(2) $q=0.5$ $\alpha=33.0$	(3) $q=0.8$ $\alpha=32.4$
BB $2q\alpha$	13.44	33	51.84
AB $(q-p)\alpha$	-20.16	0	19.44
AA $-2p\alpha$	-53.76	-33	-12.96

57 （8.6）（1）除表明不同窝序（第一、第二，等）之间的方差已经从母猪内方差即从 V_{Ew} 中消除了之外，窝序的均方是无关的。重复力由组内相关给出，$r = \sigma_b^2/(\sigma_b^2 + \sigma_w^2)$，其中 σ_b^2 是母猪间组分，σ_w^2 是母猪内组分。母猪内均方是 σ_w^2。母猪间均方由 $\sigma_w^2 + 10\sigma_b^2$ 组成。（10 是因为每一头母猪有 10 窝仔猪。）因此：

$$\sigma_b^2 = (25.56 - 3.23)/10 = 2.233$$

$$r = 2.233/(2.233 + 3.23) = 0.409$$

由组内相关进行的重复力估计假定在连续的窝中方差没有差异。

（2）以一窝以上的平均值为基础的度量降低了 V_P（见[8.14]），相对的减少量由［8.15］给出。V_A 不变。当只有一窝被使用时，遗传力是 V_A/V_P，当 n 窝被使用时，它是 $V_A/V_{P(n)}$。因此，具有 n 窝的遗传力相对于具有 1 窝的遗传力是 $V_P/V_{P(n)}$，它简单地就是［8.15］的倒数。

n	$V_{P(n)}/V_P$	h_n^2/h^2
2	0.704	1.42
3	0.606	1.65
4	0.557	1.80

采用 2 窝将提高遗传力 42%，采用 4 窝提高遗传力 80%。

58 （10.2）如果对雄性和雌性之间方差的差异进行校正，那么可以消除儿子和女儿之间的不一致性。回归和其标准误要乘以的校正因子是：对女儿—父亲，$\sigma_m/\sigma_f = 2.5/2.3 = 1.087$；对儿子—母亲，$\sigma_f/\sigma_m = 0.920$。通过作这种校正，然后加倍回归和其标准误而获得的遗传力估计值如下。

	父亲	母亲
儿子	0.646 ± 0.116	0.835 ± 0.105
女儿	0.633 ± 0.096	0.840 ± 0.096
平均值	0.64 ± 0.075	0.84 ± 0.071

来自母亲的估计值远远高于来自父亲的估计值，这可归因于母体效应。因此对中亲值的回归没有用，对父亲的回归提供了最可靠的估计值。没有理由偏爱儿子或女儿，因此我们取平均值为 $h^2 = 0.64 \pm 0.075$。获得的标准误为 $\frac{1}{2}\sqrt{0.116^2 + 0.096^2}$。

现在考察对母亲的回归。通过加倍回归值，像表中一样，我们还要加倍由母体效应引

起的协方差。因此$\frac{1}{2}$（0.84 - 0.64）= 0.10 估计了母亲和她的孩子们的环境协方差，表示为 V_P 的比例。

59 （11.6）平均而言，必须从每一头母绵羊中选择 2 只羊羔以代替亲本。在一个繁殖季节之后，没有足够的羊羔来提供更新，因此亲本必须饲养至少 2 个季节。在 2 个配种季节之后，每一只母羊平均有 2.4 只羊羔；从 2.4 只中选择 2 只（p = 83%）使得 i = 0.305（来自附录表 A，因为数目很大）。世代间隔是它们选择的后裔出生时双亲的平均年龄，在这个例子中它是 2.5 年。于是每年的选择强度是 0.122。下表展示了双亲被淘汰年龄的每一年的计算过程，直到第 7 年。当亲本繁殖了 5 个季节为 6 岁时，每年的选择强度最大。因此，亲本应当在其第 5 个繁殖季节之后被淘汰。由于下面的原因，每一年选择的比例是从 3 个中选择 1 个。每一年 5 对中有一对被更新，在那一年平均产生 1.2 只羊羔。因此更新它们所需要的总羊羔的比例是：

$$\frac{1}{5}\times\frac{2}{1.2}=\frac{1}{3}$$

或者，当到了更新的时候，每一对已平均产生总数为 6 的羊羔。因此要更新每一对，必须选择 2/6 = 1/3。

淘汰时亲本年龄	世代间隔（L）	每只母羊的总羊羔（N）	$p=2/N$（%）	i	i/L
3	2.5	2.4	83	0.305	0.122
4	3.0	3.6	56	0.704	0.235
5	3.5	4.8	42	0.931	0.266
6	4.0	6.0	33	1.097	0.274
7	4.5	7.2	28	1.202	0.267

60 （19.3）从附录表 A 中获取选择强度，给出：

雄性：i = 1.755

雌性：i = 1.400

平均：i = 1.5775

直接反应根据［11.3］来预测，相关反应根据［19.6］来预测。来自习题 19.1 的所需数据是：

$h_C^2 = 0.52$　　$h_F^2 = 0.40$　　$r_A = 0.71$

σ_C = 111g　　σ_F = 248g

代入［19.6］，给出 $h_C h_F r_A = 0.324$。

增重的反应

直接：　$R = 1.5775 \times 0.52 \times 111 = 91.1$g

相关：　$CR = 1.5775 \times 0.324 \times 111 = 56.7$g

饲料消耗的反应

直接：$R = 1.5775 \times 0.40 \times 248 = 156.5$g

相关：$CR = 1.5775 \times 0.324 \times 248 = 126.8$g

61 （1.7）亲本的基因频率已知，应当用来计算期望值。*so* 的频率是 8/40 = 0.2。根据［1.2］，后代中哈迪—温伯格基因型频率是 0.04、0.32、0.64。将这些频率乘以计数的总数 1 441 给出期望数（最接近的整数）为：

so/*so*	*so*/*cn*	*cn*/*cn*
58	461	922

有过量的纯合体，有相应不足的杂合体。可能的原因是：一些雌亲在它们被放入小瓶之前已和它们自己种群中雄亲交配引起的选型交配。差异是高度显著的，$\chi^2 = 125$。这个 χ^2 有两个自由度，因为后裔的观察数目没有用来估计基因频率，唯一的约束是期望数必须加起来等于观察总数。

62 （2.6）$q_0^2 = 1/2\,500 = 0.0004$，$q_0 = 0.02$。根据［2.12］，起始突变率是 $u_0 = sq_0^2$，对 $s = 1$，$u_0 = q_0^2$。由突变引起的基因频率变化可由［2.3］给出，由选择引起的基因频率变化可近似地由［2.8］给出。让 u_1 代表新的突变率，净改变为：

$$\Delta q = u_1(1-q_0) - sq_0^2(1-q_0)$$
$$= (u_1 - sq_0^2)(1-q_0)$$

代入 $sq_0^2 = u_0$

$$\Delta q = (u_1 - u_0)(1-q_0)$$

令 $u_1 = 2u_0$，且代入 $u_0 = q_0^2$

$$\Delta q = q_0^2(1-q_0)$$
$$= 0.0004 \times 0.98$$
$$= 0.000392$$
$$q_1 = q_0 + \Delta q$$
$$= 0.020392$$
$$q_1^2 = 0.0004158$$ 或 1/2 405。

发生率在其起始水平上增加 4%。每百万出生婴儿中将有 16 个额外的病例。

63 （3.7）（1）样本中基因频率是 0.6 和 0.4，从中可知杂合体的哈迪—温伯格频率是 0.48。观察的频率超过了期望值，表明某种形式的选择不利于一种或两种纯合体。（2）对来自这么几个亲本的样本，雄性和雌性亲本之间基因频率的随机差异变得很重要。这里 $N = 8$，根据［3.16］，杂合体的期望频率是：

$$H = 0.48(1 + 1/16) = 0.51$$

这接近于观察频率，因此选择的证据消失了。

64 （5.7）采用［5.1］，通过系谱分析，或者采用［5.3］的规则，通过共祖率可以求解这个问题。但直接从纯合体频率中获得近交系数更为简单。这借助定义给出近交系数，因为基因来自高度近交的系，因此所有的纯合体对血统同一的等位基因一定是纯合的。称近交系为 A 和 B。所有的 F_2 个体以相等的比例产生 A 和 B 配子，F_1 也是一样。因此与 F_1 回交产生 $\frac{1}{4}$AA + $\frac{1}{4}$BB，和 A 回交产生 $\frac{1}{2}$AA。因此两者产生 50% 的纯合体，两种后代的近交系数是 0.5。

65 （20.2）让 Y 表示智商得分，W 表示由家系大小度量的适合度。根据［20.5*b*］，智商得分的变化可预测为：

$$CR_Y = r_A h_Y h_W \sigma_Y \sigma_W$$

取 h_W^2 为 0.1 给出：

$$CR_Y = 0.11 \times \sqrt{0.6 \times 0.1} \times 15.4 \times 2.3$$
$$= 0.95$$

取 h_W^2 为 0.2 给出 $CR_Y = 1.35$。预测的变化是每世代智商大约增加一个点。

根据［20.4］，相关选择差为表型协方差：

$$S'_Y = \mathrm{cov}_P = r_P \sigma_Y \sigma_W$$
$$= 0.11 \times 15.4 \times 2.3$$
$$= 3.9 \text{ 个智商点}$$

66 （7.7）育种值和显性离差给在表 7-9 中。在解习题 7.3 中，我们发现 $a = 10$，$d = 50$，$q = 0.4$。（q 是所要求的减效等位基因的频率。）首先必须从［7.5］中找出基因替代的平均效应。这就是：

$$\alpha = 10 + 50\ (0.4 - 0.6) = 0$$

因此所有的育种值都是 0。（其理由是群体处于其最大的平均值，如果性状是适合度，这将是它的平衡值。）代入 p，q 和 d 给出显性离差为：

A_1A_1	A_1A_2	A_2A_2
−16	+24	−36

通过观察来检查群体中个体的平均显性离差为 0。

67 （12.1）计算展示在表中。下面的解释可能是需要的。

$$\sigma_A = h\ \sigma_P = \sqrt{0.2} \times 2.10 = 0.94$$

在极限时的总反应（一个方向），$R = 23.55 - 11.90 = 11.65$。

半存留期：这可以推导如下。总反应的一半是 5.8。给定 $S = 2.25$ 和 $h^2 = 0.2$，根据［11.2］每世代的反应是 $2.25 \times 0.2 = 0.45$。如果反应是线性的，那么产生反应为 5.8 所需要的世代数是 $5.8/0.45 = 13$。这估计了半存留期。

根据［12.2］，注意 $i\sigma_P$ = 选择差 = 2.25，理论上的最大反应 $R_{(\max)} = 2 \times 33 \times 0.2 \times 2.25 = 29.70$。

假定所有基因是加性的，那么理论上最大的半存留期是 $1.4N_e = 46$。

一个方向的反应，	H/C	= 23.55/11.90	= 2.0
	R/σ_A	= 11.65/0.94	= 12.4
	R/σ_P	= 11.65/2.10	= 5.5
$R_{(\max)}/\sigma_P$		= 29.70/2.10	= 14.1
N_e		（给定）	= 33
持续期		（给定）	= 34
半存留期		（推导）	= 13
半存留期/N_e		= 13/33	= 0.39
观察值/最大值：	反应	= 11.65/29.70	= 0.39
	半存留期	= 13/46	= 0.28
［12.1］“位点”数		$n = \frac{(17.6)^2}{8 \times (0.94)^2}$	= 44（77）
$2a/\sigma_P$		$= 2h\sqrt{2/n}$	= 0.19（0.14）

不知道较低的选择极限就不能计算位点数目及其标准化效应（n 和 $2a/\sigma_P$）。人们可能猜想向下和向上选择产生相等的反应，使得总范围 $R=23.3$；括号中的数字是以这个假设为基础的。可是，这个假设使得处于较低极限的生长不可能小，即 $11.90-11.65=0.25$g。一个更合理的猜想是向下和向上选择将产生相同比例的变化；向上选择使生长加倍，因此向下选择使生长减半，使得较低极限为 5.95 以及总范围 $R=17.6$。首先进入的数字是以这个假设为基础的。可是，记住基因数目的估计值取决于不大可能真实的几个其他的假设。

68 （10.3）因为我们操作的是相关而不是协方差，所以组分全部是总表型方差 V_P 的比例。父系半同胞给出的估计值为：

$V_A/V_P=4\times0.140=0.560$。

就标准误而论，这与来自习题 10.2 孩子对父亲回归的估计值是一致的。从母系和父系半同胞相关之间的差异中估计出对同一母亲的孩子共同的环境方差为：

$V_{Ec}/V_P=0.257-0.140=0.117$。

同一母亲的孩子之间的环境相似性预计是习题 10.2 中发现的孩子和母亲之间环境相似性的结果。这两个协方差几乎是相同的，虽然这不是期望值的一个必要部分。全同胞相关估计为 $(\frac{1}{2}V_A+\frac{1}{4}V_D+V_{Ec})/V_P$，因此：

$$\frac{1}{4}V_D/V_P=0.406-2\ (0.140)\ -0.117=0.009$$

$$V_D/V_P=0.036$$

不求出其标准误，也容易看出这与 0 没有显著差异。剩下的组分，即同胞关系内的环境方差可从与总数为 1 的差异中获得。以百分率表示的完全的分剖是：

$$V_A\ +V_D\ +V_{Ec}+V_{Ew}=V_P$$

$$56\ +4\ \ +12\ +28\ \ =100$$

69 （14.1）由［14.4］给出平均值的变化为 $-2F\sum dpq$。在一代自体受精后，$F=0.5$（表 5-2），因此由 4 个位点引起的近交衰退是 $\sum dpq$，其中 d 是杂合体值和纯合体中间值之间的差异（图 7-1）。纯合体中间值在下面给出为与“AA”值的差异，是数据表中“aa”值的一半。

位点	纯合体中间值	d	pq	dpq
(1)	−10	0	0.25	0
(2)	−15	+20	0.25	5
(3)	−15	−5	0.16	−0.8
(4)	−30	+30	0.09	2.7
			$\sum dpq=$	6.9

由这些位点引起的产量将减少 6.9g。

70 （19.4）根据［11.7］，现实遗传力如下：

$h^2_C=186/574=0.32$

$h^2_F=525/1312=0.40$

根据［19.7］，遗传相关是：

$$r^2_A=\frac{120}{186}\times\frac{412}{525}=0.5063$$

$$r_A=\pm0.71$$

因为直接的和相关的反应处于同一方向，r_A 的符号一定是正的，因此 $r_A = +0.71$。

71 （1.8）不期望具有哈迪—温伯格频率，因为雄性和雌性亲本中基因频率不同。*so* 的基因频率，在雄性中是 16/20 = 0.8，在雌性中是 4/20 = 0.2。将这些作为配子频率放到表1-8中，给出后代中基因型频率为：

so/*so*	*so*/*cn*	*cn*/*cn*
0.16	0.68	0.16

对基因频率为0.5，哈迪—温伯格期望值是0.25，0.50，0.25。注意过量的杂合体是由雄性和雌性亲本中不相等的基因频率引起的。

72 （2.7）让 q 表示突变等位基因的频率，因此 $q_0 = 1$。让 m 表示具有 $q_m = 0$ 的迁入者的比例（$m = 0.01$）。于是［2.1］给出：

$$q_1 = mq_m + (1 - m)q_0$$
$$= (0.01 \times 0) + (0.99 \times 1) = 0.99$$
$$q_2 = (0.01 \times 0) + (0.99 \times 0.99) = (0.99)^2$$
$$q_{10} = (0.99)^{10} = 0.9044$$

野生型等位基因的频率 $= p = 1 - 0.9044 = 0.0956$

在最后的迁入者繁殖之后，基因型将处于哈迪—温伯格比例。因此

（1）野生型果蝇的频率 $= p^2 = (0.0956)^2 = 0.0091$

（2）突变表型的频率 $= 1 - p^2 = 0.9909$

杂合体的频率 $= 2pq = 0.1729$

突变体中杂合体的频率 $= 0.1729/0.9909 = 0.1745$

73 （3.8）从杂合体的期望频率开始，给出：

$$H = 2pq + \frac{1}{2}\overline{D^2} = 2pq + \frac{1}{2}\sigma_D^2$$

其中，D 是雄性和雌性亲本之间基因频率的差异。假定有 M 个雄性和 F 个雌性亲本，抽取了 $2M$ 和 $2F$ 个基因。基因频率的二项抽样方差，在雄性亲本中是 $pq/(2M)$，在雌性亲本中是 $pq/(2F)$，其中 p 和 q 是整个群体的总基因频率。基因频率差异的抽样方差是：

$$\sigma_D^2 = pq/(2M) + pq/(2F)$$

因此，修改的等式是：

$$H = 2pq + pq\left(\frac{1}{4M} + \frac{1}{4F}\right)$$
$$= 2pq\left[1 + \left(\frac{1}{8M} + \frac{1}{8F}\right)\right]$$

74 （5.8）考虑仅由表亲结婚组成的群体的一部分。这是具有 $F = 1/16$ 的一个细分群体（见习题5.1）。孩子的危险率是纯合体的频率，在表3-1中给出，为 $q^2 + pqF$。称这种危险率为 Q。为了评估 Q，我们需要基因频率 q。因为整个群体中大部分婚姻是非血亲的，我们可以取 q 近似地为发生率

	(1) 囊性纤维变性	(2) PKU
发生率，$I\ (=q^2)$	4×10^{-4}	90.91×10^{-6}
$q = \sqrt{I}$	2×10^{-2}	9.53×10^{-3}
Fpq	12.25×10^{-4}	590×10^{-6}
$Q = q^2 + Fpq$	16.25×10^{-4}	681×10^{-6}
$R = Q/I$	4.1	7.5

的平方根（这样做有一种小误差，因为有一些表亲结婚，基因型不完全处于哈迪一温伯格比例；见习题 5.9）。最后，表亲结婚的孩子相对于整个群体的危险率是 $R = Q/I$，其中 I 是群体发生率。

囊性纤维变性的危险率增至 4 倍，PKU 的危险率增至 7.5 倍，但从绝对值看这种危险率仍然很小。

75 （15.1）基础群体参数：

$V_P = 4.0$；　$V_G = V_A = 0.52 \times 4.0 = 2.08$；　$V_E = 4.0 - 2.08 = 1.92$

这是缓慢近交，因此我们使用表 15-1 的栏目（1）。对 $F = 0.5$，这给出

	遗传的	环境的	表型的
系间	$2FV_G = 2.08$	0	2.08
系内	$(1-F)\ V_G = 1.04$	1.92	2.96
总的	$(1+F)\ V_G = 3.12$	1.92	5.04

遗传力：　（1）系内 1.04/2.96 = 0.35

（2）总的 3.12/5.04 = 0.62

76 （7.8）发现的两个基因有关的值是：

基因 b：　$a = d = 2.5$；$q = 0.5$

基因 c^e：　$a = d = 28.5$；$q = 0.2$

首先从［7.5］中获得每一个基因替代的平均效应：

基因 b：　$\alpha = 2.5 + 2.5\ (0.5 - 0.5) = 2.5$

基因 c^e：　$\alpha = 28.5 + 28.5\ (0.2 - 0.8) = 11.4$

我们关注的基因型对应于表 7-9 中的 A_2A_2。对每一个位点分别代入上面的值，给出下面的结果。

	育种值 $= -2p\alpha$	显性离差 $= -2p^2d$
基因 b：	$-2 \times 0.5 \times 2.5 = -2.50$	$-2 \times (0.5)^2 \times 2.5 = -1.25$
基因 c^e：	$-2 \times 0.8 \times 11.4 = -18.24$	$-2 \times (0.8)^2 \times 28.5 = -36.48$
两个基因：	−20.74	−37.73

将单个位点的育种值加起来给出联合基因型的育种值；对显性离差也是类似的。计算的育种值是与群体平均值的离差，发现群体平均值是 91.47 个颗粒。因此，采用绝对单位，育种值是 91.47 − 20.74 = 70.73 个颗粒。

77 （13.1）可根据［13.4］和［13.5］来计算遗传力，也给出在表 13-4 中。从表 13-4 的右边栏中计算了相对反应。

注意，当 $t = r$ 时，很容易从［13.4］和［13.5］中看出 $h_f^2 = h^2 = h_w^2$。在这些条件下，个体选择给出了个体和家系平均值的最好的加权。还要注意到使家系选择或家系内选择优于个体选择的条件。

	h_f^2 ［13.4］	h_w^2 ［13.5］	R_f/R	R_w/R
(1)	0.265	0.077	0.93	0.72
(2)	0.390	0.077	1.06	0.74
(3)	0.100	0.100	0.79	0.61
(4)	0.147	0.500	0.68	0.97
(5)	0.136	0.500	0.62	1.05

78 （10.4）均方的期望是：

$$B = \sigma_W^2 + 2\ \sigma_B^2$$

$W = \sigma_W^2$

从中有：

$\sigma_B^2 = (B - W)/2$

$\sigma_W^2 = W$

组内相关是：

$$t = \sigma_B^2/(\sigma_B^2 + \sigma_W^2)$$

$$= \frac{\frac{1}{2}(B - W)}{\frac{1}{2}(B - W) + W}$$

$$= \frac{B - W}{B + W}$$

79 （14.2）根据［7.3］，群体平均值是 $M = \sum a\ (p - q) + 2\sum dpq$。这是与多重纯合体中值的离差。全部纯合的有利等位基因的平均值将是 $\sum a$，也是与多重纯合体中值的离差。因此增加将是：

位点	a	aq
(1)	10	5
(2)	15	7.5
(3)	15	3
(4)	30	3
$\sum aq$	=	18.5

$$\sum a - M$$

$$= \sum a - \sum a(p - q) - 2\sum dpq$$

$$= \sum[a(2q)] - 2\sum dpq$$

$$= 2(\sum aq - \sum dpq)$$

我们需要知道每一个位点的 a 值。这是纯合体值之间差异的一半。在习题 14.1 中获得的 $\sum dpq$ 值为 6.9。

增加 = 2（18.5 - 6.9）= 23.2g。

80 （19.5）解答来自［19.9］，对体重的选择产生了窝产仔数的相关反应。根据附录表 A，对应于中选比例的选择强度是

	选择	
	窝产仔数（X）	体重（Y）
对雌性的 i	1.271	1.271
对雄性的 i	0	1.755
i（平均）	0.6355	1.513

代入［19.9］给出：

$$\frac{CR}{R} = 0.43 \times \frac{1.513}{0.6355} \times \sqrt{\frac{0.35}{0.22}}$$

$$= 1.29$$

主要是因为能对雄性选择，所以预计选择体重比选择窝产仔数的有效性高 29%。

81 （1.9）通过小数目等位基因的“试验和误差”可容易地证明它是正确的。（对都具有相同频率的 3 个等位基因计算杂合体的频率；然后用不相等的频率重新计算。）一种简单的一般证明来自对等位基因频率方差的考察。对相等的频率，方差为 0。让 n 表示等位基因的数目，q 表示任一等位基因的频率。那么 $\sum q = 1$，$(\sum q)^2 = 1$。纯合体的频率是 $\sum q^2$，当这是最小值时，杂合体频率则最大。q 的方差由下式给出：

$$\sigma_q^2 = \frac{1}{n}[\sum q^2 - \frac{(\sum q)^2}{n}] = \frac{1}{n}[\sum q^2 - \frac{1}{n}]$$

重新整理导致：

$$\sum q^2 = n\sigma_q^2 + \frac{1}{n}$$

因此，对 n 的任何值，当 $\sigma_q^2 = 0$ 时，纯合体的频率最小。于是所有等位基因具有相等的频率 $1/n$。

将 $\sigma_q^2 = 0$ 代入上面的等式给出纯合体的频率为 $1/n$。因此杂合体的频率是$1 - (1/n)$。

注意上面第一个等式可采用下面有用的形式来书写：

$$\sigma_q^2 = \frac{\sum q^2}{n} - [\frac{\sum q}{n}]^2 = \overline{q^2} - (\overline{q})^2$$

即方差 = （平方的平均值） -（平均值的平方）。这在以后各章中用到。

82 （2.8）因为两个群体达到了相同的中间的基因频率，所以选择一定有利于杂合体。这是预料之中的，因为“杂合体” $so+/+cn$ 是野生型，而两个纯合体是突变型。不利于两个纯合体的选择系数的相对大小可由［2.18］得到：

$$\frac{s_{(so)}}{s_{(cn)}} = \frac{p_{(cn)}}{q_{(so)}} = \frac{0.65}{0.35} = 1.86$$

83 （4.1）由［4.4b］给出有效群体大小。然后通过［4.1］获得 ΔF，由［3.12］获得近交系数。

雌性数目	10	10	10	10
雄性数目	10	5	2	1
N_e	20	13.33	6.67	3.636
ΔF	0.025	0.0375	0.075	0.1375
$F(t=10)$	0.224	0.318	0.541	0.772

84 （5.9）从表 3-1 的右手项看，在非近交个体中发生率是 q^2，在近交个体中发生率是 $q^2(1-F) + qF$。因此总的发生率是：

$$I = (1-y)q^2 + y[q^2(1-F) + qF]$$

通过打开方括号，这很容易简化成：

$$(1-yF)q^2 + yFq = I$$

如果 y 和 F 已知，这就可以解出 q。

注意 yF 是整个群体的平均近交系数，通过用 yF 来代替 F，可从表 3-1 中直接获得总的发生率的表达式。

假定没有偏离哈迪—温伯格比例的其他原因，这个等式可用来得出习题 5.8 的精确解。表亲结婚的频率变化很大，但在很多群体中大约是 1%。

85 （15.2）这是快速近交，因此我们应当使用表 15-1 的栏目（2） $F = 0.5$，$f = 0.594$（来自表 5-2）。这给出了下表左边的组分。可是，我们需要从 4 个个体中估计的系群观察平均值的方差，它是 $\sigma_b^2 + \frac{1}{4}\sigma_w^2$（表 13-3）。这给出下表右边的值。

	组分		观察平均值的方差	
	遗传的	表型的	遗传的	表型的
系群间 σ_b^2	2.47	2.47	2.63	3.11
系群内 σ_w^2	0.65	2.57	—	—

观察到的系群平均值的遗传力是 2.63/3.11 = 0.85。根据附录表 A，p = 5% 的选择强度是 i = 2.063。期望反应由［13.2］给出，系群等价于家系。它是

$R = 2.063 \times 0.85 \times \sqrt{3.11} = 3.1$ 根刚毛。

(如果作为缓慢近交来求解，R = 2.9 根刚毛：没有很大的差异。)

86 (7.9) 为了获得群体平均值，我们需要 4 种基因型的频率，这显示在表 (i) 中。将例 7.7 中的每一种基因型值乘以它的频率，然后相加给出群体平均值 = 1.112。将值转换成如表 (ii) 所示的与平均值的离差。现在这些与平均值的离差将简单地被称为值。下一步，我们必须分别考察每一个位点，并获得在这个群体中它的两个基因型每一个的平均值。这些给在表 (ii) 的边缘；例如，C - 基因型的平均值是 0.6 (0.328) + 0.4 (-0.342) = +0.060。现在得到像表 (iii) 那样的联合基因型的加性期望值。如果两个单个位点的基因型值简单地加在一起，这些就是联合基因型将具有的值。例如，B - C - 基因型的期望值是 0.228 (B -) + 0.060 (C -) = 0.288。最后，每一种基因型的互作离差是表 (ii) 中的观察值与表 (iii) 中的加性期望值之间的差异。这些给在表 (iv) 中。例如，B - C - 的互作离差是 0.328 - 0.288 = +0.04。为了检验，看看平均互作离差是否为 0。

(i) 频率

	频率	B - 0.6	bb 0.4
C -	0.8	0.48	0.32
c^rc^r	0.2	0.12	0.08

(ii) 观察的与平均值的离差

	频率	B - 0.6	bb 0.4	平均值
C -	0.8	0.328	-0.342	+0.060
c^rc^r	0.2	-0.172	-0.342	-0.240
平均值		+0.228	-0.342	0.000

(iii) 加性期望值

	B -	bb
C -	+0.288	-0.282
c^rc^r	-0.012	-0.582

(iv) 互作离差

	B -	bb
C -	+0.04	-0.06
c^rc^r	-0.16	+0.24

互作离差还可以直接从如下的基因型值的表中计算出来。边缘的平均值与上面表 (ii) 中的一样获得。

	B -	bb	平均值
C -	1.44	0.77	1.172
c^rc^r	0.94	0.77	0.872
平均值	1.34	0.77	1.112

于是，选取 B - C - 基因型作为一个例子，在计算互作离差过程中紧接上面的步骤可汇总为

$$(1.44 - M) - [(1.34 - M) + (1.172 - M)]$$

其中 M 是群体平均值 = 1.112。于是这种基因型的互作离差变成：

$$1.44 - 1.34 - 1.172 + 1.112 = 0.04$$

87 (13.2) 相对于个体选择 R 的反应可从表 13-4 的右边的表达式中获得。(1) 和 (2) 是家系选择；(3) 是同胞选择；(4) 是家系内选择。

(1) $$R_f/R = \frac{1 + (4 \times 0.25)}{\sqrt{5\,[1 + (4 \times 0.10)]}} = 0.76$$

(2) $R_f/R=\dfrac{1+(4\times0.5)}{\sqrt{5\ [1+(4\times0.36)]}}=0.86$

(3) $R_s/R=\dfrac{5\times0.5}{\sqrt{5\ [1+(4\times0.36)]}}=0.72$

(4) $R_w/R=(1-0.5)\sqrt{\dfrac{4}{5\ (1-0.36)}}=0.56$

88 (10.5) 这说明了在解释双生子数据中的一些困难。可以采用 3 种有不同偏差的方法来估计遗传力，如下表所示。所有的估计都是有偏差的，但大小不同，由显性和由上位组分引起的偏差没有显示在表中。(2) 和 (3) 是由共同环境造成偏差的。10 岁以下的孩子作出了有理性地良好判断。遗传决定程度由 (1) 近似地估计为 52%。存在由共同环境引起的相似性，通过从 (2) 中减去 (1) 或等价地从 (3) 中减去 (2)，可近似地估计它为 $V_{Ec}=12\%$。10～15 岁的孩子没有作出良好判断。它表明 V_{Ec} 对他们的相似性没有贡献，因为 (2) 和 (3) 低于 (1)。

		10 岁以下	10～15 岁
(1) 2 (MZ－DZ)	$=(V_A+\frac{3}{2}V_D)/V_P$	0.52	0.98
(2) MZ	$=(V_A+V_D+V_{Ec})/V_P$	0.64	0.91
(3) 2DZ	$=(V_A+\frac{1}{2}V_D+2V_{Ec})/V_P$	0.76	0.84

89 (14.3) 实验为 $F=0$ 的平均值给出了两个值 8.1 和 8.5。解答将以这两个值的平均值即 8.3 为基础。当 $F=1$ 时，假定 D_M 和 D_L 分别表示在母亲中和在窝仔畜中由近交引起的衰退。然后根据例 14.2 中表的第三行有：

$$0.5\,D_M=8.3-6.2$$

$$D_M=4.2$$

根据表的第二行：

$$0.5D_L+0.375\,D_M=8.3-5.7$$

$$D_L=2.05$$

母亲中近交引起的衰退大约是窝仔畜中近交引起的衰退的两倍。$F=1$ 时总的衰退是：

$$D_M+D_L=4.2+2.05=6.25$$

平均窝产仔数将是 $8.3-6.25=2.05$ 头幼畜/窝。

90 (19.6) 首先求出仅对雌性选择的期望反应，在解答习题 19.5 过程中没计算它。我们可以假定从大数目中选择出 25%，因此选择强度可从附录表 A 中获取。根据 [11.3]，反应是：

$$R=\frac{1}{2}\times1.271\times0.22\times\sigma_P=0.140\,\sigma_P$$

对联合选择，我们通过 [11.3] 来计算预计来自直接选择雌性窝产仔数的反应以及通过 [19.6] 来计算来自对雄性体重选择的相关反应，然后把两个期望反应加在一起。这等于计算中选雌性和雄性窝产仔数的平均育种值。单位自始至终都是窝产仔数的表型标准差。对雌性的选择强度必须从附录表 B 中获取，因为全部都是从小样本中选择的，即从 4

个中选择 1 个。预测反应是：

来自雌性：$R=\frac{1}{2}\times 1.029\times 0.22$ $=0.113\ \sigma_P$

来自雄性：$CR=\frac{1}{2}\times 1.271\times\sqrt{0.35\times 0.22}\times 0.43$ $=0.076\ \sigma_P$

来自两性：联合反应 $=0.189\ \sigma_P$

相对于只对雌性选择的期望反应，联合反应是：

$$\frac{0.189}{0.140}=1.35$$

第三种方法比第一种方法要好 35%。

91 (1.10) 求解来自表 1-16 和等式 [1.5]。AA bb 是由两个 Ab 型重组配子结合而产生的，因此我们必须找出它的频率 s。(Ab 对应于表 1-16 中的 A_1B_2)。我们需要知道下面的数量。(i) 基因频率；在两个位点上这都是 0.5。(ii) Ab 配子的平衡频率 $\hat{s}$；这是 $\hat{s}=p_Aq_B=0.25$。(iii) 在任何重组发生之前，0 世代中不平衡的度量 D_0，(iv) 2 世代中的不平衡 D_2。从 Ab 配子频率中计算的任何世代的不平衡是 $-D=s-\hat{s}$。0 世代中，$s=0$，因此 $D_0=\hat{s}=0.25$。D_2 从 [1.5] 中获得；2 世代后代是 2 个世代重组的产物，因此 $t=2$，$D_2=D_0\ (1-c)^2$。对问题 (1) 中的自由重组，$c=0.5$，$D_2=0.25\ (0.5)^2=0.0625$。问题 (2) 中 $c=0.2$，$D_2=0.25\ (0.8)^2=0.16$。接下来我们返回到上面给出的等式 $-D=s-\hat{s}$。将这个书写为 $s=\hat{s}-D_2$ 并代入获得的 $\hat{s}$ 和 D_2 值，我们得到 2 世代，

(1) $s=0.25-0.0625=0.1875$，　　(2) $s=0.25-0.16=0.09$。

最后，由这些配子产生的后代中 AA bb 的频率是 s^2。因此答案是 (1) 0.0352 和 (2) 0.0081。

92 (2.9) 假定 R 表示频率为 p 的抗性基因，S 表示易感的等位基因；假定 s_1 表示不利于 RR 的选择系数，给定 $s_1=0.63$。为了找出因为中毒而死去的大鼠的比例，我们必须找出不利于 SS 的选择系数 s_2。根据 [2.18]，

$$\frac{s_2}{s_1}=\frac{p}{1-p}$$

$$s_2=0.63\times\frac{0.34}{0.66}=0.32$$

死亡的比例是：

RR	SS
s_1p^2	$+s_2\ (1-p)^2$
$=0.63\ (0.34)^2$	$+0.32\ (0.66)^2$
$=0.07$	$+0.14$
$=0.21$	

所有大鼠的 21% 死亡。这通过 [2.21] 可以更直接地得到，它给出总死亡，即负荷为 $L=0.63\times 0.34=0.21$。

93 (4.2) 将 $N_m=N_f/d$ 代入 [4.4a] 给出：

$$\frac{1}{N_e}=\frac{d}{4N_f}+\frac{1}{4N_f}$$

$$= \frac{d+1}{4N_f}$$

$$N_e = \frac{4N_f}{d+1}$$

94 （18.1）所需要的 x 和 i 值可用内插法从附录表 A 中获得。相关 t 可从近似公式［18.1］中计算。r 的值来自表 9-3。遗传力通过［18.2］来估计。以第一级亲缘为例，

$$t = (3.090 - 2.061)/3.367 = 0.306$$

$$h^2 = 2t = 0.61$$

	p%	x	i	t	r	h^2%
群体	0.1	3.090	3.367	—	—	—
亲缘						
第 1 级	1.97	2.061	—	0.306	$\frac{1}{2}$	61
第 2 级	0.34	2.706	—	0.114	$\frac{1}{4}$	46
第 3 级	0.19	2.895	—	0.058	$\frac{1}{8}$	46

来自第一级亲属的估计值可能是最精确的，即具有最小的标准误，因为患病的亲属数最多，并因为 t 的标准误乘以 2 而不是乘以 4 或 8。另一方面，第一级亲缘由于子宫内母体效应可能具有某种环境相关。最可靠的估计值可能来自第二级亲缘。近似标准误可从等式［18.2］下面的段落中给出的 t 的抽样方差的表达式中计算出来，标准误是抽样方差的平方根。将 t 的标准误乘以 $\frac{1}{r}$ 给出 h^2 3 种估计值的标准误分别为 4%、10% 和 26%。来自第一级和第二级亲缘的估计值彼此之间没有显著差异，来自第三级亲缘的估计值与 0 没有显著差异。

95 （15.3）第一世代是全同胞，使得 $F = 0.25$。假定没有进一步的近交，并且非加性遗传方差可以忽略。于是重新整理［15.1］导致：

$$h_0^2 = \frac{h_t^2}{1 - F\ (1 - h_t^2)}$$

代入 $h_t^2 = 0.34$ 给出基础群体的遗传力为：

$$h_0^2 = \frac{0.34}{1 - \ (0.25 \times 0.66)} = 0.41$$

96 （16.1）三元杂交的预测产量是两个单元杂交的平均值。因此我们必须寻找包含 3 个品种在内的最好的两个单元杂交。按价值的顺序，单元杂交产量是：

杂交	AE	AC	BC	AB	DE	BE
产量	31.8	22.8	16.5	14.1	13.1	12.4

最好的两个单元杂交包含了 3 个品种，因此是适合的。在两个杂交中出现的品种将必须用于第二次杂交，因此杂交是（C × E）× A，它的预测产量是 $\frac{1}{2}$（AC + AE）= 27.3。

四元杂交的预测产量是 4 个单元杂交的平均值。因此我们必须寻找具有 4 个品种的最好的 4 个单元杂交，并且 4 个品种中的每一个都包含在两个单元杂交中。这些是 AE，AC，BC 和 BE。杂交将为（A × B）×（C × E），它的预测产量是 $\frac{1}{4}$（AC + AE + BC + BE）= 20.9。

97 （13.3）所需要的指数是针对个体和同胞平均值的。表 13-3 中观察到的家系平均值方差的加性和表型组分系数分别用 k 和 K 表示。

$$k=\frac{1+(4\times 0.5)}{5}=0.600$$

$$K=\frac{1+(4\times 0.36)}{5}=0.488$$

指数中加权因子是：

$$b_1=\frac{h^2(0.400)}{0.512};\quad b_2=\frac{h^2(0.112)}{0.250},$$

指数是：

$$I=h^2(0.781)P_1+h^2(0.448)P_2,$$

其中 P_1 是个体增重，P_2 是家系均值，两者都是与群体平均值的离差。

为应用的方便，指数按比例重新调整为：

$$I'=P_1+0.574P_2$$

这个按比例重新调整的指数也可以从［13.8］来计算。

注意如果指数的单位是增重的单位，那么按比例重新调整的指数中的 P_2 以及未调整指数中的 P_1 和 P_2 一定是与群体平均值的离差。

98 （10.6）我们首先必须计算如表 10-6 所示的观察组分。代入 $d=3$，$k=10$，雄性的均方是：

$$3.894=\sigma_W^2+10\,\sigma_D^2+30\,\sigma_S^2$$

$$2.198=\sigma_W^2+10\,\sigma_D^2$$

$$1.125=\sigma_W^2$$

根据这些等式以及对应雌性的等式，观察组分和相关如下。

	雄 性	雌 性
组分		
公畜间 σ_S^2	0.05653	0.0800
母畜间 σ_D^2	0.1073	0.1168
母畜内 σ_W^2	1.125	0.893
总的 σ_T^2	1.28883	1.0898
相关		
半同胞 σ_S^2/σ_T^2	0.0439	0.0734
全同胞 $(\sigma_S^2+\sigma_D^2)/\sigma_T^2$	0.1271	0.1806

观察和成因组分之间的关系给在表 10-7 中。成因组分的估计值如下。

没有母系半同胞，V_{Ec}不能从$\frac{1}{4}V_D$ 中分离出来，V_{Ew}也不能从$\frac{3}{4}V_D$ 中分离出来。

		刚毛单位		总数的百分数	
		雄性	雌性	雄性	雌性
V_A	$=4\sigma_S^2$	0.2261	0.3200	17.5	29.4
$\frac{1}{4}V_D+V_{Ec}$	$=\sigma_D^2-\sigma_S^2$	0.0508	0.0368	3.9	3.4
$\frac{3}{4}V_D+V_{Esc}$	$=\sigma_W^2-2\sigma_S^2$	1.0119	0.7330	78.5	67.3
V_P	$=\sigma_T^2$	1.2888	1.0898	99.9	100.1

99 （14.4）预测的衰退是 $0.56D_M+0.64D_L$。取习题 14.3 中对非近交平均值 8.3 计算的 D_M 和 D_L 值，这给出衰退为：

$(0.56\times4.2)+(0.64\times2.05)=3.66$

图 14-2（a）中的平均值，从曲线中可以读出，$F=0$ 时近似为 7.6，最后一个世代中近似为 4.2，这给出衰退为每窝 3.4 头幼畜。这与预测值吻合得相当好。

100 （19.7）假定下标 1 表示生长（G），下标 2 表示饲料消耗（F）。来自［19.11］的指数方程是：

$$b_1P_{11}+b_2P_{12}=A_{11}$$

$$b_1P_{21}+b_2P_{22}=A_{21}$$

将给定的参数估计值代入到［19.12］：

$$P_{11}=(1.11)^2=1.2321$$

$$P_{22}=(2.48)^2=6.1504$$

$$P_{12}=P_{21}=0.83\times1.11\times2.48=2.2848$$

$$A_{11}=0.52\times(1.11)^2=0.6407$$

$$A_{21}=0.71\times\sqrt{0.52\times0.40}\times1.11\times2.48=0.8914$$

将这些值代入指数方程：

$$1.2321b_1+2.2848b_2=0.6407 \quad (1)$$

$$2.2848b_1+6.1504b_2=0.8914 \quad (2)$$

(1) $\times6.1504$：$7.5779b_1+14.05b_2=3.9406$ (3)

(2) $\times2.2848$：$5.2203b_1+14.05b_2=2.0367$ (4)

(3) − (4)：$2.3576b_1=1.9039$

$$b_1=0.8076$$

将 b_1 代入（1）：

$$b_2=-0.1551$$

指数是：

$$I=0.808G-0.155F$$

或重新调整，采用下列形式给 G 单位权重，

$$I'=G+WF$$

其中 $W=b_2/b_1$，它是：

$$I'=G-0.192F$$

101 （1.11）现在单个位点基因型在 0 世代后代中不处于哈迪—温伯格比例，因为在

雄性和雌性亲本中基因频率不同。这影响到1世代的不平衡，因此这种不平衡必须首先解出。重组配子只由 AB/ab 基因型产生。这里所有0世代后裔都是这种基因型的，因此 s_1（Ab 配子的频率）是 $\frac{1}{2}c$。（在习题1.10中，0世代的后裔只有一半是 AB/ab，因此 s_1 是一半的大。）计算的其余过程如右表。

	(1) $c=0.5$	(2) $c=0.2$
$s_1=\frac{1}{2}c$	0.25	0.10
平衡 $\hat{s}$	0.25	0.25
$D_1=\hat{s}-s_1$	0	0.15
由 [1.5]，$D_2=D_1(1-c)$	0	0.12
$s_2=\hat{s}-D_2$	0.25	0.13
s_2^2（AA bb 的频率）	0.0625	0.0169

“杂交”要比习题1.10中品系“混合”后的不平衡小。没有连锁时，在1世代中两位点平衡频率就达到了，它与基因型频率为1/16的典型的两因子杂交 F_2 相符合。

102 （2.10）亲本（两个基因）之间的基因型频率和突变型的基因频率 q_0：

AA	Aa	aa	q_0
0.6	0	0.4	0.4

基因（a）配子频率：

	A	a	总数
	0.6	0.4×0.5	0.8
除以0.8：	0.75	0.25	1.0

这些配子之间随机结合给出后代中基因型频率：

AA	Aa	aa
0.5625	0.3760	0.0625

根据 [1.1]，后代中观察的频率是 $q_1=0.25$。这和配子中的相同。因此以 q_1 为基础的哈迪—温伯格期望值恰好与观察值相同。单个的后裔不能提供选择的证据。

基因（b）：突变体的配子频率是 $q_0=0.4$。对随机交配，后代的基因型频率是：

	BB	Bb	bb	总数
合子中	0.36	0.48	0.16	1.00
存活者中	0.36	0.48	0.08	0.92
观察值	0.391	0.522	0.087	1.000

根据 [1.1]，观察的基因频率：$q_1=0.348$，这给出哈迪—温伯格期望值为：

0.425	0.454	0.121

观察到的频率有过量的杂合体和不足的两个纯合体，表明忽视了有利于杂合体的选择。单个的后裔仅告诉我们哈迪—温伯格期望值的条件没有全部满足。基因（a）的 Δq 较大，因为在亲本间没有杂合体，纯合体中所有突变基因都因选择而暴露。对基因（b），杂合体中很多突变基因都躲避了选择。

103 （4.3）记住500个繁殖对子意味着 $N=1\,000$。

根据 [4.6]，$N_e=253$；

根据 [4.1]，$\Delta F=0.20\%$。

104 （18.2）所需要的 x 和 i 值来自附录表A并给在下面。根据 [18.1] 计算了相关 t。相关乘以2给出遗传力，因为 $r=\frac{1}{2}$。这两个估计值吻合得相当好。“重复的出生”以

同样的方式作为亲属来处理，给出相关，它是重复力。经计算 t 值为（1.812 - 1.282）/ 2.208 = 0.24。

	p%	x	i	t	h^2%
群体	3.5	1.812	2.208	—	—
母亲	4.6	1.685	—	0.0575	12
女儿	4.8	1.665	—	0.0666	13
重复出生	10.0	1.282	—	0.24	—

105 （15.4）发生的突变有$\frac{1}{4}$将固定在一个同胞交配的系群中，其中每一个世代有 4 个配子。因此发现的突变数是每一个配子的突变率。分离亚系群所需的世代数是 21 + 29 = 50，并有 27 个性状。因此每一世代每一个配子每一性状的总突变率是：

$$\frac{5}{50 \times 27} = 3.7 \times 10^{-3}$$

这是总突变率。如果每一个位点的突变率是 u，并且如果每一个性状的 n 个位点中的任一位点受突变影响，那么

$$3.7 \times 10^{-3} = nu$$

至于什么是 u 和 n 值的最可靠联合，则可以采取我们的选择。例如

u	n
3.7×10^{-3}	1
3.7×10^{-4}	10
3.7×10^{-5}	100

106 （16.2）假定 A、B 和 AB 代表两个系群或品种以及单交或 F_1 的表现。

世代	杂交的基因型	后代的期望表现
1	AA × BB	AB
2	AB × AA	$(AB + A)$ /2
3	AB × BB AA × BB	$(3AB + B)$ /4
4	3AB × AA BB × AA	$(5AB + 3A)$ /8

107 （13.4）我们使用重新调整的指数，其中 P_1 是个体增重，P_2 是家系平均值，它必须表达为与群体平均值的离差；b_2 是 P_2 的权重因子。对 $n = 5$ 的家系，习题 13.3 给出 $b_2 = 0.574$。对 $n = 8$，$k = 0.5625$ 和 $K = 0.4400$，给出 $b_2 = 0.636$。个体的指数值是：

A：$1.6 + [0.574 \times (-0.2)] = 1.485$

B：$1.5 + (0.574 \times 0.1) = 1.557$

C：$1.5 + (0.636 \times 0.1) = 1.564$

D：$1.3 + (0.636 \times 0.2) = 1.427$

价值顺序是 C、B、A、D。

（注意：P_2 必须是与群体平均值的离差，因为它的权重（b_2）在所有个体中不是相同的。P_1 不必是一个离差，因为它的权重是相同的。有一个其平均值与群体平均值或零相等的指数是明智的。如果指数中所有的度量都表达为离差，那么指数本身就是一个离差，它的平均值是 0。如果一个度量值采用实际单位，那么所有其他的度量值必须是离差，于是指数的平均值等于群体平均值。）

108 （10.7）首先考察 $t=0.75$，$h^2=0.5$ 的特定情形。设 $V_P=100$。于是 $V_A=50$，$\mathrm{cov}_{FS}=75$。

$$\mathrm{cov}_{FS}=\frac{1}{2}V_A+\frac{1}{4}V_D+V_{Ec}=75$$

因此

$$\frac{1}{4}V_D+V_{Ec}=75-25=50$$

现在表型方差可以写为：

$$V_P=V_A+(\frac{1}{4}V_D+V_{Ec})+\frac{3}{4}V_D+V_{Ew}=100$$

因此

$$\frac{3}{4}V_D+V_{Ew}=100-50-50=0$$

因此 h^2 不能大于 0.5，因为无论是 V_D 还是 V_{Ew}都不可能小于 0。

为了得出一般结论，考察全同胞家系内方差（见表 10-7），它是：

$$\frac{1}{2}V_A+\frac{3}{4}V_D+V_{Ew}=V_P-\mathrm{cov}_{FS}=V_P\ (1-t)$$

$$\frac{1}{2}V_A=V_P\ (1-t)\ -\frac{3}{4}V_D-V_{Ew}$$

因此当 V_D 和 V_{Ew}为 0 时，V_A 最大。这样在最大值处的 h^2 值为：

$$\frac{1}{2}h^2=1-t$$

$$h^2=2\ (1-t)$$

如果全同胞相关是 0.8，则最大可能的遗传力是：

$$h^2=2\times0.2=0.4$$

具有 $V_{Ew}=0$ 的性状是极不可能的，尤其是如果 V_{Ec}不为 0，因此在实践中 h^2 的上限将大大低于 $2\ (1-t)$。

109 （14.5）我们将计算每一世代需要的雌性平均选择差。$N_e=40$ 时，由［4.1］知近交率 $\Delta F=0.0125$。根据［3.12］，第 30 世代的近交系数是 $F=0.3143$。对这样的缓慢近交，将母亲和窝仔畜的近交系数视为相同，是足够好的近似。来自母亲和窝仔畜在一起的衰退将取自习题 14.3 的解答，$F=1$ 时为 6.25。因此 $F=0.3143$ 的预期衰退是 $0.3143\times6.25=1.964$，每世代的衰退速度是 $1.964/30=0.065$。抵消选择将需要每世代产生 0.065 的反应。将 $h^2=0.22$ 代入［11.2］，这使得 $S=0.065/0.22=0.30$。由于对雄性没有选择，雌性的选择差平均来说必须是每世代每窝 0.6 只小鼠。这意味着，例如，如果在每一世代内 5 只雌性中有 1 只雌性按照选择目标被更新，则它们必将来自平均每窝比它们的群体平均值高出 $0.6\times5=3$ 只小鼠的窝仔畜中。

110 （19.8）指数选择的预测反应由［19.17］给出，但我们首先需要根据［19.18］

来找出指数值的方差，为此，我们需要未调整的值 $b_1 = 0.8076$ 和 $b_2 = -0.1551$。根据[19.18]

$$\sigma_I^2 = b_1 A_{11} + b_2 A_{21}$$
$$= (0.8076 \times 0.6407) + (-0.1551 \times 0.8914)$$
$$= 0.3792$$
$$\sigma_I = 0.6158$$

根据[19.17]，指数选择的反应是：

$$R_I = 1.5775 \times 0.6158$$
$$= 0.971$$

在习题 19.7 中指数的计算单位是 100g。因此预测的改进是每世代 97g。

在习题 19.3 中计算了单独对生长做选择的反应。以 100g 为单位，它是 $R = 0.911$。因此指数选择的相对效率是：

$$\frac{R_I}{R} = \frac{0.971}{0.911} = 1.066$$

指数预期要好 6.6%。在这种情形中，次级性状的使用产生了极少的利益。

111 (11.7) 现在对雄性和雌性选择强度和世代长度都不同。假定在成熟时雄性和雌性数目是相等的。每只母羊每个季节的羔羊平均数像习题 11.6 表述的一样为 1.2 只。因此每只雄性亲本每个季节的雄性羔羊数目为 $\frac{1}{2} \times 1.2 \times 10 = 6$。雄性繁殖 2 个季节，因此每个雄性平均有 12 个雄性后代，从中必须选择出 1 个雄性后代，给出 $i = 1.840$。为雌性更新，对雌性的选择强度是 $i = 0.704$，与习题 11.6 中计算的相同。对雄性世代长度是 2.5 年，对雌性是 3 年。因此每年的平均选择强度是：

$$\frac{\frac{1}{2}(1.840 + 0.704)}{\frac{1}{2}(2.5 + 3)} = 0.463$$

相对于习题 11.6 中最优方法的效率是 0.463/0.274 = 1.69，它将要好 69%。

为什么淘汰亲本有一个最适年龄的理由是，L 以相等的步幅增加，但 i 却以减小的步幅增加。最适年龄较低的理由是对雄性的选择强度较大。淘汰两性的最适年龄取决于平均的 i 和平均的 L，而不取决于每一性别中分别的 i 与 L 的比率。

112 (2.11) 对雄性和雌性的选择效应不同，必须分别求解。异配性别将指定为雄性。让 q 代表隐性基因的起始频率。基因型和它们的起始频率以及存活者之间和下一世代的基因频率列在下面。存活雌性之间的频率见[2.9]的推导。

两性存活者之间总基因频率是：

$q_1 = \frac{2q}{3(1+q)}$ (见[1.4])。

下一世代的合子中，基因频率不再与两个性别中的相同，但总基因频率不再进一步改变。因此我们能够采用上面 q_1 的表达式作为一个世代选择之后的总基因频率。于是基因频率的变化是：

$$\Delta q = \frac{2q}{3(1+q)} - q$$

这可简化成：

$$\Delta q = -\frac{q\ (1+3q)}{3\ (1+q)}$$

	雄 性		雌 性		
	A	a	AA	Aa	aa
	p	q	p^2	$2pq$	q^2
基因频率					
存活者中	0		$q/\ (1+q)$		
下一世代中	$q/\ (1+q)$		$\frac{1}{2}q\ (1+q)$		

113 （4.4）作为一个理想化群体繁殖自体可育的品种，对它来说，$N_e = N$；$\Delta F = \frac{1}{2N_e} = 2.50\%$。根据［4.2$b$］，对自体可育品种，$\Delta F = 2.44\%$。自体受精的排除产生极小的差异，对较大的 N 引起的差异甚至更小。

114 （18.3）这种形式的分析等于对“0，1”尺度进行操作。如果指定给个体的值是 0 和 1 而不是 1 和 2，这不会引起差异。每一个半同胞家系有一个平均值（值为 0，1），该平均值就是它产生双生子的成员比例。因此半同胞的协方差就是半同胞家系之间比例的方差。于是，在 0，1 尺度上，遗传力是 $4 \times 0.0058 = 2.32\%$。通过［18.4］，这可以转换成趋向尺度。我们需要群体发生率和对应的 i，从习题 18.2 中可知，它们是 $p = 0.035$，$i = 2.208$。于是趋向遗传力是：

$$h^2 = \frac{2.32 \times 0.965}{(2.208)^2 \times 0.035}\% = 13\%$$

这与习题 18.2 中通过趋向分析获得的估计值非常紧密地一致。

115 （15.5）杂种间的方差组分 σ_X^2 直接来自［15.4］。总方差与基础群体的方差相同，因此杂种内组分 σ_W^2 是与 V_P 的差异。杂种间的方差组分 σ_X^2 是杂种真实平均值的方差。观察均值的方差 σ_M^2 是 $\sigma_X^2 + \frac{1}{20}\sigma_W^2$（见表 13-3）。

	F	杂种间 $FV_A + F^2V_D = \sigma_X^2$	杂种内 $\sigma_W^2 = V_P - \sigma_X^2$	观察平均值 $\sigma_M^2 = \sigma_X^2 + \frac{1}{20}\sigma_W^2$
(1)	0.5	432 + 47 = 479	5 319	745
(2)	1.0	864 + 188 = 1 052	4 746	1 289

116 （16.3）假定 A，B 和 C 代表纯种表现，AB，AC，BC 代表单交表现。

世代	杂交基因型	预期的后代表现
1	AA × BB	AB
2	AB × CC	$(AC + BC)/2$
3	AC × AA	$(2AC + AB + A)/4$
	BC × AA	
4	2AC × BB	
	AB × BB	$(5AB + 2BC + B)/8$
	AA × BB	
5	5AB × CC	
	2BC × CC	$(9BC + 5AC + 2C)/16$
	BB × CC	

117 （13.5）为计算效率，我们需要未调整指数的系数。在习题 13.3 中发现这些值是：

$$b_1 = 0.781h^2;\ b_2 = 0.448h^2$$

不必要知道 h^2，因为当与个体选择进行比较时，h^2 将会消掉。

接下来我们需要由［13.12］给出指数方差。在这里，$A_{11} = \sigma_A^2$；A_{21}是个体与其家系平均值的加性协方差，等于观察的家系平均值的加性方差，它为 $k\sigma_A^2$。于是代入［13.12］给出：

$$\sigma_I^2 = (b_1 + b_2 k)h^2\ \sigma_A^2$$

$$= 1.050h^2\ \sigma_A^2$$

指数选择的期望反应由[13.15]给出，为：

$$R_I = i\sigma_I = i(1.025)h\sigma_A$$

这与［11.4］中个体选择的期望反应 R 是直接可比的，并且只要选择强度 i 相同，就有

$$R_I/R = 1.025$$

指数选择将比个体选择好 2.5%。获利是小的，因为 t 与 r 没有很大差异。

118 （10.8）所有 3 种回归和全同胞相关显著不同于 0，但有大的标准误，方差的分剖只是试验性的。后代对中亲值的回归估计了遗传力，没有选型交配引起的偏差，给出 $h^2 = 0.82$。用表 10-14 的等式（14）来校正由选型交配（$r = 0.33$）造成的对一个亲本的回归偏差，分别给出 $h^2 = 2b/(1 + r) = 0.71$ 和 0.72。这些与来自中亲值回归的估计值是一致的。为了处理选型交配对同胞相关的影响，我们需要知道育种值的相关 m。不知道这个值，我们可以取 $m = r$ 作为一个近似值。于是根据表 10-14 的等式（15），并且取 $h^2 = 0.82$，则全同胞关于育种值的相关是 $t = 0.41 \times 1.33 = 0.55$。观察到的相关高于这个值，虽然不显著，表明对全同胞有一些共同的环境方差，或一些显性方差，共计为总方差的 $0.71 - 0.55 = 16\%$。

119 （14.6）假定 $\overline{P}$ 表示两个亲本的平均产量，并假定 F_1，F_2，F_3 代表这些世代的产量。那么预测的产量是：

$$F_2 = \frac{1}{2}(\overline{P} + F_1)$$

$$F_3 = \frac{1}{2}(\overline{P} + F_2)$$

给出

	观察值		预测值	
杂交	$\overline{P}$	F_1	F_2	F_3
(1)	1.40	1.41	1.405	1.40
(2)	1.08	1.42	1.25	1.165

通过考察杂种优势 H，这些期望值的理由可能变得更清楚。根据定义，$H_{F_1} = F_1 - \overline{P}$。根据［14.10］，$H_{F_2} = \frac{1}{2} H_{F_1}$。$F_2$ 的预期产量是：

$$\overline{P} + H_{F_2} = \overline{P} + \frac{1}{2}(F_1 - \overline{P}) = \frac{1}{2}(\overline{P} + F_1)$$

每一世代自体受精使杂合体的频率减少一半（表 5-2），因此 F_3 代预期的杂种优势是 F_2 代的一半。因此 F_3 的预测产量是：

$$\overline{P} + \frac{1}{2}(F_2 - \overline{P}) = \frac{1}{2}(\overline{P} + F_2)。$$

120 （19.9）相关反应由［19.19］给出，但我们首先需要得到性状 2 与指数的加性遗传协方差。性状 1 和指数的协方差由［19.20］给出。性状 2 和指数的协方差，它给在例 19.5 中，是：

$$\mathrm{cov}_{2I} = b_2 A_{22} + b_1 A_{21}$$

在习题 19.7 中 A_{22} 没有计算。它是 $0.40 \times (2.48)^2 = 2.4602$。

$$\mathrm{cov}_{2I} = (-0.1551 \times 2.4602) + (0.8076 \times 0.8914)$$
$$= 0.3383$$

为了从［19.19］中得到反应，我们需要 $i = 1.5775$ 和 $\sigma_I = 0.6158$，两者均来自习题 19.8。于是饲料消耗的相关反应是：

$$CR_2 = 1.5775 \times \frac{0.3383}{0.6158}$$
$$= 0.87 \text{（以 100g 为单位）}$$

饲料消耗预计每世代将增加 87g。

121 （11.8）假定 n 代表选择基于的窝数，假定 t 代表以年为单位的世代长度。假定 R 和 R_n 分别代表当选择以 1 和 n 窝为基础时的反应。反应可通过［11.3］，$R = ih^2\sigma_P$ 来比较，但通过［11.4］，$R = ih\sigma_A$ 来比较更简单，因为 σ_A 不随着窝数而改变。我们不必评估实际反应，因为最适宜的 n 值将是使下列比率达到最大值的数目。

$$\frac{R_n/t_n}{R/t} = \frac{i_n h_n}{i h} \frac{t}{t_n}$$

计算展示在下面。注意到每一窝中雌性的预期数目是 4，因此必须选择的母猪比例对 $n = 1$ 是 1/4，对 $n = 2$ 是 1/8，等等。在习题 8.6 中对重复力 0.409 计算了 h_n^2/h^2。i 值可从附录表 A 中查到。

最适宜的窝数是 2。

n	t	$p\%$	i	h_n^2/h^2	i_n/i	h_n/h	t/t_n	$\frac{R_n/t_n}{R/t}$
1	1	25	1.271	—	—	—	—	—
2	1.5	12.5	1.647	1.42	1.296	1.192	0.667	1.03
3	2	8.33	1.840	1.65	1.448	1.285	0.500	0.93

122 (2.12) 作为一种近似，假定雌性纯合体有害基因的频率低至可以忽略。于是通过选择淘汰这个基因仅发生在雄性中。基因的 1/3 在雄性中，因此每一世代中通过选择可淘汰基因的比例是 $sq/3$，其中 s 是选择系数。平衡时，这些为突变所取代，因此 $sq/3 = u$，$q = 3u/s$。

肌肉营养不良的数据与 $s = 1$ 的这个期望值吻合得很好，没有不利于杂合雌性的选择。忽略雄性—雌性的差异，基因频率，是雄性的发生率，正如所期望的那样，它几乎恰好是估计突变率的 3 倍。

123 (4.5) 假定 k = 使用的后代数，即给定的家系大小。亲本的实际数目，包括不育的对子，是 $N = 16$。为了从 [4.7] 中获得 N_e，我们需要计算 k 的方差。平均值是 $\bar{k} = 2$，离差 $(k - \bar{k})$ 是：-2，-1，-1，0，0，$+1$，$+1$，$+2$，从中可知 $\sum (k - \bar{k})^2 = 12$，$V_k = 12/8 = 1.5$。(除以 8 而不是 7，因为这是统计意义上的整个“群体”，不是一个样本。) 于是由 [4.7]，$N_e = 64/3.5 = 18.3$。尽管没有使家系大小相等，但 N_e 仍然比 N 稍大一些。

124 (18.4) 对 x 和 i 的符号需要小心。这里 x 将表示作为与阈值离差的群体平均值，因此发生率低于 50% 时，x 是负的。感染个体的平均值 i_A 是与群体平均值的离差，是正的。我们还将需要未感染个体的平均值 i_N，这是负的。附录表 A 提供了发生率为 p 的 i_A 值，从中可得到 i_N 为 $i_N = -i_A p/(1-p)$。两个世代必须分别求解，因为第一次选择引起的发生率变化改变了第二次选择的选择差。

首先将解释降低发生率的选择，因为它较简单。求解过程显示在下面。因为 50% 以上的个体是正常的，所有中选的两性个体将都是正常的，以标准差为单位的选择差将是 $S = i_N$。反应是 $R = h^2 S$，新的平均趋向是 $x_1 = x_0 + R$。为了得到新的发生率，我们必须在附录表 A 中找出对应于 x_1 的 p 值。给出的值是通过内插法获得的，但如果取最接近的制成表的 p 值，则几乎不产生差异。对第二次选择，使用 p_1 和 x_1 代替 p_0 和 x_0，重复上面的计算。预测发生率将减少到 17%。

降低发生率的选择

第一次选择		第二次选择	
p_0	23%	p_1	19.6%
x_0	−0.739	x_1	−0.857
i_A	1.320	i_A	1.411
$S = i_N$	−0.394	$S = i_N$	−0.344
$R = 0.3S$	−0.118	$R = 0.3S$	−0.103
$x_1 = x_0 + R$	−0.857	$x_2 = x_1 + R$	−0.960
p_1	19.6%	p_2	16.9%

除选择差更复杂以外，下面显示的增加发生率的选择的计算与以前的相同。因为中选雄性的比例低于发生率，因此所有中选雄性都是被感染的，雄性的选择差是 $S_{♂} = i_A$。可是，因为被感染的低于50%，所以中选的雌性必须包括一些正常个体。在第一次选择中，被感染的中选雌性的比例是 0.23/0.5 = 0.46，正常的比例为 0.54。因此对雌性的选择差是 $S_{♀} = 0.46i_A + 0.54i_N$。或者，就发生率而言，$S_{♀} = 2pi_A + (1-2p)\ i_N$。两性合并选择将是 $\frac{1}{2}(S_{♂} + S_{♀})$。其余的计算不必进一步解释。在第二次选择后，发生率将增加到 41%。对增加发生率的选择大得多的反应是由较大的选择差引起的。

增加发生率的选择

第一次选择		第二次选择	
p_0	23%	p_1	31.5%
x_0	-0.739	x_1	-0.482
i_A	1.320	i_A	1.128
i_N	-0.394	i_N	-0.519
$S_{♂} = i_A$	1.320	$S_{♂}$	1.128
$S_{♀} = 2pi_A + (1-2p)\ i_N$	0.3944	$S_{♀}$	0.5186
$S = \frac{1}{2}(S_{♂} + S_{♀})$	0.8572	S	0.8233
$R = 0.3S$	0.257	R	0.247
x_1	-0.482	x_2	-0.235
p_1	31.5%	p_2	40.7%

125 (15.6) 假定 X 表示一个杂种的真实平均值，M 表示观察平均值。从附录表 B 可知，从 50 个中选择 1 个给出 $i = 2.249$。因此最好的杂种与所有杂种平均值的离差是 $i\sigma_M$，其中 σ_M 是观察平均值的标准差。在习题 15.5 中计算过 σ_M^2。所有杂种的期望平均值是起始群体平均值，为 308。将 308 加到 $i\sigma_M$ 上给出最好杂种的预测观察平均值。

	F	σ_X^2	σ_M^2	σ_M	$i\sigma_M$	预测的 M
(1)	0.5	479	745	27.3	61.4	369.4
(2)	1.0	1 052	1 289	35.9	80.7	388.7

未来表现的预测在第 8 章有关重复力的部分做过解释。重复杂种的期望平均值，即未来表现，是杂种的真实平均值。为了预测真实平均值，我们需要知道真实平均值对观察平均值的回归。观察平均值是 $M = X + E$，其中 E 是由抽样误差引起的离差。现在

$$\mathrm{cov}_{XM} = \mathrm{cov}_{X(X+E)} = \mathrm{cov}_{XX} + \mathrm{cov}_{XE}$$

$\mathrm{cov}_{XX} = \sigma_X^2$，$\mathrm{cov}_{XE} = 0$，因为 X 和 E 不相关。因此 $\mathrm{cov}_{XM} = \sigma_X^2$。所需要的回归是 $b_{XM} = \mathrm{cov}_{XM}/\sigma_M^2 = \sigma_X^2/\sigma_M^2$。其余的计算如下。

	$b_{XM}=\sigma_X^2/\sigma_M^2$	最好杂种的预测真实平均值
(1)	0.643	308 + 0.643(61.4) = 347.5
(2)	0.816	308 + 0.816(80.7) = 373.9

126 (16.4) 在每一种情形中，解答表中右边栏代表后代中每一种基因型的比例。考察纯合体的比例是最简单的，相对于没有纯合体的第一世代，它是近交系数 F。于是，相对于单交的杂种优势大小是 $H_t = 1 - F_t$。

世代 (t)	两个系群		二个系群	
	F_t	H_t	F_t	H_t
1	0	1	0	1
2	1/2	0.5	0	1
3	1/4	0.75	1/4	0.75
4	3/8	0.625	1/8	0.875
5	5/16	0.6875	2/16	0.875
很多	1/3	0.667	1/7	0.857

如果连续世代继续下去，在几个循环之后将发现在两系群的情形中杂种优势固定在 2/3，在二系群的情形中固定在 6/7。相对于第一次杂交的最后近交水平的一般公式，是 $1/(2^n - 1)$，其中 n 是系群数目。

127 (13.6) 所需要的指数是对母亲和半同胞姐妹平均值的。指数方程是：

$$b_2 P_{22} + b_3 P_{23} = A_{21}$$

$$b_2 P_{32} + b_3 P_{33} = A_{31}$$

下标 1 指要选择的个体，它没有被度量；2 指母亲，3 指半同胞姐妹平均值。$P_{22} = \sigma^2$（即表型方差）；$P_{23} = P_{32} = 0$；$P_{33} = K\sigma^2$，其中 $K = [1 + (n-1)t]/n$，与以前一样。A_{21} 是个体与其母亲的加性协方差 $= r_2\,\sigma_A^2 = r_2 h^2\,\sigma^2$，其中 $r_2 = \frac{1}{2}$。A_{31} 是个体与不包括个体在内的它的半同胞姐妹平均值的加性协方差；这是与同胞选择有关的协方差，这不受半同胞数目的影响，其理由在紧接着等式 [9.2] 的段落中作过解释。因此 $A_{31} = r_3 h^2\,\sigma^2$，其中 $r_3 = \frac{1}{4}$。将这些值代入指数方程中，给出：

$$b_2\,\sigma^2 + 0 = r_2 h^2 \sigma^2$$

$$0 + b_3 K \sigma^2 = r_3 h^2 \sigma^2$$

据此

$$b_2 = r_2 h^2 = \frac{1}{2} h^2$$

$$b_3 = r_3 h^2 / K = \frac{1}{4} h^2 / K$$

现在估计 K。因为半同胞姐妹之间没有环境相似性，$t - \frac{1}{4} h^2 = 0.0875$。对于 $n = 10$，这给

出：

$$K=\frac{1+(9\times 0.0875)}{10}=0.17875$$

代入 K 和 $h^2=0.35$，给出未调整的指数为：

$$I=0.175P_2+0.490P_3$$

128 （10.9）表 10-14 中的等式提供了解答。因为配偶的选择纯粹是表型的，$m=rh^2$。于是等式（15）是 $t=\frac{1}{2}h^2(1+rh^2)$，其中 h^2 是选型交配的群体遗传力。重新整理给出 $\frac{1}{2}rh^4+\frac{1}{2}h^2-t=0$，它的解为：

$$h^2=\left[-\frac{1}{2}\pm\sqrt{\left(\frac{1}{4}+2rt\right)}\right]/r$$

代入 r 和 t 的值给出唯一一个可能的解，它是 $h^2=0.50$。如果它是随机交配的，则同一群体的遗传力 h_0^2 由等式（9）给出。代入 $h^2=0.5$ 和 $m=rh^2=0.2$ 给出：

$$h_0^2=0.5(0.8/0.9)=0.44$$

129 （14.7）从［14.10］中可看出 F_1 和 F_2 之间的差异是杂种优势的一半，即 $F_1-F_2=\frac{1}{2}H$，从中可知 $H=-0.24$。但 F_1 和 F_2 之间的差异很小且不显著，因此推导出的这个小的杂种优势与 0 没有显著差异。实际上，亲本品种的平均值是 17.88 ± 0.24 和 15.00 ± 0.19，这使得 $\overline{P}=16.44\pm0.20$，观察到的杂种优势是 $H=-0.72\pm0.28$。

130 （19.10）来自［19.15］的指数方程是：

$$b_1P_{11}+b_2P_{12}=a_1A_{11}+a_2A_{12}$$

$$b_1P_{21}+b_2P_{22}=a_1A_{21}+a_2A_{22}$$

左边，是表型参数，与习题 19.7 相同。对于右边，我们需要 $A_{11}=0.6407$ 和 $A_{12}=A_{21}=0.8914$，两者也来自习题 19.7。像习题 19.9 那样计算，$A_{22}=2.4602$。以 100g 为单位的经济值已经运用于方程中其余参数，因此 $a_1=8$ 和 $a_2=-2$。将这些值代入，指数方程是：

$$1.2321b_1+2.2848b_2=(8\times0.6407)+(-2\times0.8914)=3.3428$$

$$2.2848b_1+6.1504b_2=(8\times0.8914)+(-2\times2.4602)=2.2108$$

像习题 19.7 那样消除 b_2 给出：

$$2.3576b_1=20.5596-5.0512$$

$$b_1=6.578$$

$$b_2=-2.084$$

用于选择的指数是：

$$I=6.578G-2.084F$$

或者，更方便地，

$$I'=G-0.317F$$

131 （11.9）我们首先必须根据［11.8］获得对基因起作用的选择系数 s。代入 $2a=0.3$，$\sigma_P=2.0$ 和 $i=1.755$ 给出 $s=0.2633$。于是一个世代选择之后的基因频率可由表 2-2 的行（1）给出。公式中基因频率 q 是选择不利的等位基因频率，因此我们必须代入 $q=0.6$。公式可以用更简单的一种形式来重写：

$$q_1 = q\left[1 - \frac{1}{2}s(1+q)\right]/(1-sq)$$
$$= 0.6[1 - 0.1316(1.6)]/(1 - 0.1580)$$
$$= 0.5625$$

对第二世代，使 q_1 代替 q，给出：

$$q_2 = 0.5245$$

因此增效等位基因的频率将是0.4755。

132 (2.13) 红色的起始基因频率 $q_0 = \sqrt{0.01} = 0.10$。(1) 根据 [2.9]，$q_2 = 0.1/1.2 = 0.0833$。红色犊牛的频率 $= q_2^2 = 0.0069$。(2) 这是难处理的，因为对公牛的选择比对母牛的多。因此，选择之后，在雄性和雌性亲本中基因频率不相同，后代的基因型不是哈迪—温伯格频率。必须分别对每一个性别计算基因频率的变化。选择前杂合体的频率是0.18。因为检测中红色犊牛不能逃避检测，所以杂合公牛的比例是 $\left(\frac{3}{4}\right)^6 = 0.178$。因此作为未检测出杂合体的所有公牛的比例是 $0.18 \times 0.178 = 0.0320$。

		公牛			母牛			
第一世代	RR	Rr	rr	总数	RR	Rr	rr	总数
选择前	0.81	0.18	0.01	1.00	0.81	0.18	0.01	1.00
选择后	0.81	0.0320	0	0.8420	0.81	0.18	0	0.99
	0.9620	0.0380	0	1.0000	0.8182	0.1818	0	1.000
	$q_{1m} = \frac{1}{2} \times 0.0380 = 0.0190$				$q_{1f} = \frac{1}{2} \times 0.1818 = 0.0909$			

现在像表1-8那样制一张配子频率表，得到后代基因型频率。

		雄性配子	
		R	r
		0.9810	0.0190
雌性	R 0.9091	0.8918	0.0173
配子	r 0.0909	0.0892	0.0017

		公牛			母牛			
第二世代	RR	Rr	rr	总数	RR	Rr	rr	总数
选择前	0.8918	0.1065	0.0017	1.0000	0.8918	0.1065	0.0017	1.0000
选择后	0.8918	0.0190	0	0.9108	0.8918	0.1065	0	0.9983
	0.9792	0.0208	0	1.0000	0.8933	0.1067	0	1.0000
	$q_{2m} = 0.0104$				$q_{2f} = 0.05335$			

后代中红色犊牛的频率是 $q_{2m}q_{2f} = 0.06\%$。

（注意：q_{1f}可以更容易地通过 [2.9] 得到，但 q_{2f}不能用这种方法得到，因为第二次选择前基因型不具有哈迪—温伯格频率。）

133 (4.6) 育种计划是具有 $N=16$ 的最小近交，因此根据 [4.9]，$N_e=31$。数据导致第 27 世代的近交系数为 $F(t=27)=0.447$。根据[3.12]

$(1-\Delta F)^{27}$ $=1-0.447=0.553$

$27\log(1-\Delta F)$ $=-0.2573$

$1-\Delta F$ $=0.9783$

ΔF $=0.0217$

$N_e=1/(2\Delta F)$ $=23.0$

134 (18.5) 这里 x 是阈值与群体平均值的离差。所需的 x 和 i 值如下。

类别	阈值	p%	x	i
N	T_1	20	−0.842	−1.400
$H+F$	T_1	80	−0.842	$+1.400\times 20/80=+0.350$
F	T_2	30	+0.524	+1.159

$p=80\%$ 的 i 值是 $p=20\%$ 的值乘以 $(1-p)/p$，正如附录表 A 在表头中所表述的那样。

(1)阈值之间的差异

$=0.524-(-0.842)=1.366\ \sigma$

$1\text{t.u.}=1.366\ \sigma;\ \sigma=0.732\text{t.u.}$

(2) 作为与较低阈值离差的群体平均值

$=+0.842\ \sigma=(0.842\times 0.732)\text{t.u.}=+0.616\text{t.u.}$

(3) $\overline{N}=-1.400\ \sigma=-1.025\text{t.u.}$

$\overline{F}=+1.159\ \sigma=+0.848\text{t.u.}$

$(\overline{H+F})=+0.350\ \sigma=+0.256\text{t.u.}$

$(\overline{H+F})$ 是高于较低阈值的所有个体的平均值，由 50/80 的 H 和 30/80 的 F 组成。因此

$$(\overline{H+F})=\frac{50}{80}\overline{H}+\frac{30}{80}\overline{F}$$

$$0.256=0.625\ \overline{H}+(0.375\times 0.848)$$

$$\overline{H}=-0.099\text{t.u.}$$

为了检验，每一个都用它的频率进行加权，看看 N，H 和 F 的平均值求和是否为 0，作为与群体平均值的平均离差。

N：$-1.025\times 0.2=-0.2050$

H：$-0.099\times 0.5=-0.0495$

F：$+0.848\times 0.3=+0.2544$

$\sum=0.000$

135 (15.7) 求解过程完全遵循例 15.2 的过程，将不进行详细地解释。品种本身(它们的产量在表的对角线上）必须排除，因为配合力指的是杂种的表现。T 和 G 值在下表的右端。G 是表左边标明的品种的一般配合力。作为例子，

$T_B = 14.1 + 16.5 + 6.2 + 12.4 = 49.2$

$$G_B = \frac{49.2}{3} - \frac{302.6}{5 \times 3} = -3.77$$

G 是与平均值的离差；为了检验，看看是否 $\sum G = 0$。

例如，杂种 A × B 的期望值是：

$\overline{X} + G_A + G_B$

$= 15.13 + 8.36 - 3.77$

$= 19.72$

因此离差（SCA + 误差）是 $14.1 - 19.72 = -5.62$。为了检验，看看离差加起来是否为 0。离差给在下表中。

	B	C	D	E	T	G
A	-5.62	+0.55	-1.25	+6.32	85.6	+8.36
B	—	+6.38	+0.18	-0.95	49.2	-3.77
C	—	—	-0.25	-6.68	56.8	-1.24
D	—	—	—	+1.32	44.5	-5.34
E	—	—	—	—	66.5	+1.99
$\sum$	—	—	—	—	302.6	0.00

$\overline{X} = 302.6/20 = 15.13$

136 （16.5）为避免混乱，用与习题 16.3 中使用的同一字母来标明品种，如下所示。在计算每一世代的期望值中，记住任一世代的公式适用于该世代中的 W 和下一世代的 N。所需要的总窝重是 $W \times N$。

品种或杂种	Y	C	D	YC	YD	CD
符号	A	B	C	AB	AC	BC
W	84	78	88	90	96	92
N	7.9	6.6	6.3	8.2	7.3	7.4

世代	基因型		期望值		
	W	N	W	N	$W \times N$
1	AB	A	90	7.9	711
2	$X_1 \times C$	AB	94	8.2	771
3	$X_2 \times A$	$X_1 \times C$	91.5	7.35	673
4	$X_3 \times B$	$X_2 \times A$	89	7.675	683
5	$X_4 \times C$	$X_3 \times B$	92.75	7.8	723

注意杂种表现波动起伏，没有一个好于世代 2 中的三元杂种。

137 （13.7）两个选择强度不能联合起来，因为指数的标准差与产量的标准差不相

同。必须从雄性和雌性的预测育种值中获得反应。首先，根据［13.12］

$$\sigma_I^2 = b_2 A_{21} + b_3 A_{31}$$
$$= b_2 r_2 h^2 \sigma^2 + b_3 r_3 h^2 \sigma^2$$
$$= (b_2 r_2 + b_3 r_3) h^2 \sigma^2$$
$$= [(0.175 \times 0.5) + (0.490 \times 0.25)] \times 0.35 \times \sigma^2$$
$$= 0.0735 \sigma^2$$
$$\sigma_I = 0.2711 \times 696$$
$$= 188.69$$

根据附录表 A，选择 5% 的公牛给出 $i = 2.063$。中选公牛的期望育种值由［13.15］给出：

公牛育种值 $= i\sigma_I = 2.063 \times 188.69 = 389.3$

选择 50% 母牛给出 $i = 0.798$，预测育种值由［11.3］给出为：

母牛育种值 $= ih^2 \sigma_P = 0.798 \times 0.35 \times 696 = 194.4$

两个育种值都是与群体平均值的离差，期望反应是两者的平均值：

$$R = \frac{1}{2}(389.3 + 194.4)$$
$$= 292\text{kg}$$

138 （10.10）在所有情形中，我们同时需要知道家系数目 N 以及每一家系后代或同胞的数目 n。这两个值中没有给出的一个值必须从 $T = 400$ 中推导出来。N 和 n 与 T 的关系如下。

（1）$T = N(n+1)$　$N = 200$，　$n = 1$。

（2）$T = N(n+2)$　$n = 3$，　$N = 80$。

（3）$T = Nn$　$n = 5$，　$N = 80$。

（4）$T = Nn$　$N = 20$，　$n = 20$。

在（1）和（2）的情形中，每一个家系被度量的个体数是后代数 n 并在（1）中加上一个亲本，在（2）中加上两个亲本。

等式［10.6］给出了（1）和（中 2）回归的抽样方差，［10.10］给出了（3）和（4）中相关的抽样方差。而且，在（1）和（4）情形中，设计是最佳的，有较短的路径可以获得遗传力的标准误。

（1）设计是最佳的，因为 $n = 1$。具有 1 个度量的亲本（$k = 1$），［10.6］简化了［10.7］中给出的 h^2 的标准误。

$$\text{标准}(h^2\text{误}) = 2/\sqrt{200} = 0.14$$

（2）$k = 2$；　$t = \frac{1}{2}h^2 = 0.3$（忽略显性）；　$N = 80$；　$n = 3$。

$$\sigma_b^2 = \frac{2[1 + (2 \times 0.3)]}{3 \times 80} = 0.0133$$
$$\sigma_b = 0.115$$

用 b 估计遗传力，因此

标准误(h^2) = σ_b = 0.115

〔注意：这里近似的不是很好。精确的公式给出 σ_b^2 = 0.0091；标准误(h^2) = 0.095。〕

(3) $t = \frac{1}{2}h^2 = 0.2$（忽略显性）；$N = 80$；$n = 5$。

$$\sigma_t^2 = \frac{2[1+(4\times 0.2)]^2(0.8)^2}{5\times 4\times 79} = 0.00262$$

σ_t = 0.051

h^2 被估计为 $2t$，因此

标准误(h^2) = 2 σ_t = 0.10

(4) $N = 20$；$n = 20$。

对半同胞这是最佳的设计，因为 $n = 4/h^2$。遗传力的抽样方差由［10.13］近似地给出。

$\sigma_{h^2}^2 = (32\times 0.2)/400 = 0.0160$

标准误(h^2) = 0.13

精确的公式［10.10］给出标准误(h^2) = 0.12。

139 （20.3）在绘图之前所需要的一切是将死亡率转换成存活率，或许需要以标准差为单位将出生重表达为与群体平均值的离差。例如

出生重（kg）	以 σ 为单位的离差	存活率（%）
1.3	-4.24	38.8
1.8	-3.25	66.7
等等		

注意曲线的平台属性：97% 的婴儿具有在 97% ~ 99% 的狭窄范围之内存活的可能性。数据没有给出认为小婴儿因为他们小而死亡的理由；因为这些小婴儿具有引起他们死亡的某种其他缺陷，所以他们可能小。

140 （19.11）计算过程遵循习题 19.8 的计算过程。根据［9.18］，指数的方差是：

$$\sigma_I^2 = (6.578\times 3.3428) + (-2.084\times 2.2108) = 17.3816$$

σ_I = 4.169

选择强度可在习题 19.3 中找到，为 i = 1.5775。因此指数选择的期望反应是：

R_I = 1.5775 × 4.169

= 6.58

经济价值的期望改良是每世代每只鸡 6.58 美分。

当仅对生长进行选择时，生长将作为直接反应而增加，产生经济收益，而饲料消耗将作为相关反应而增加，产生经济损失。反应在习题 19.3 中计算过了。转换成 100g 单位，生长的直接反应是 + 0.911，饲料消耗的相关反应是 + 1.268。因此经济收益是：

来自生长： $0.911 \times 8 = 7.29$ 美分

来自饲料消耗： $1.268 \times (-2) = -2.54$ 美分

经济净收益 4.75 美分

因此改良经济价值的指数的相对效率是 $6.58/4.75 = 1.39$。指数要好 39%。

141 (21.1)亲本品系之间的差异是整个第 3 号染色体的效应;区间 1 的差异是区间内任何 QTL 的纯合效应。差异如下:

	腹部刚毛	腹侧刚毛
整个染色体	13.7	4.6
区间 1	2.4	2.8
占整个染色体的%	17.5%	60.9%

为了检验每一种差异的显著性,计算它的标准误,然后用 t 检验,结果如下。

	腹部刚毛	腹侧刚毛
$\sigma_d^2 = \frac{SS_H + SS_L}{n_H + n_L - 2} \times \frac{n_H + n_L}{n_H n_L}$	$\frac{651.8}{42} \times \frac{44}{403}$	$\frac{28.6}{42} \times \frac{44}{403}$
	$= 1.694$	$= 0.074$
σ_d	1.30	0.272
$t = d/\sigma_d$	1.85	10.3
d.f. $= n_H + n_L - 2$	42	42
	$P \sim 0.07$	$P < 0.001$

证实存在影响腹侧刚毛的一个 QTL,但对腹部刚毛的影响在 0.05 平准上不显著,虽然近于显著(在论文中采用的一种更灵敏的测验证明腹部刚毛效应是高度显著的)。标记类型内腹部刚毛数很高的方差是由位于染色体剩余部分上的 QTL 的分离所引起。腹侧刚毛数低得多的方差可能归因于处于分散态的仅影响腹侧刚毛的 QTL 等位基因,正如从相关反应中预测的那样。

参 考 文 献

方括号中的数字代表引用该文献的正文章序。

1 **Allard, R.W.**, **Jain, S.K.** and **Workman, P.L.** (1968) The genetics of inbreeding populations. *Adv. Genet.*, **14**, 55 ~ 131. [5]

2 **Allard, R.W.**, **Kahler, A.L.** and **Weir, B.S.** (1972) The effect of selection on esterase allozymes in a barley population. *Genetics*, **72**, 489 ~ 503. [5]

3 **Allen, J.A.** (1975) Further evidence for apostatic selection by wild passerine birds: 9:1 experiments. *Heredity*, **36**, 173 ~ 180. [2]

4 **Allison, A.C.** (1954) Notes on sickle-cell polymorphism. *Ann. Hum. Genet.*, **19**, 39 ~ 57. [2]

5 **Allison, A.C.** (1956) The sickle-cell and haemoglobin *C* genes in some African populations. *Ann. Hum. Genet.*, **21**, 67 ~ 89. [2]

6 **Aquadro, C.F.** and **Begun, D.J.** (1993) Evidence for and implications of genetic hitchhiking in the *Drosophila* genome. pp.159 ~ 178 in Takahata, N. and Clark, A.G. (eds), *Mechanisms of Molecular Evolution*. Sinauer, Sunderland, Mass., USA. [2]

7 **Arnold, S.J.** and **Wade, M.J.** (1984*a*) On the measurement of natural and sexual selection. Theory. *Evolution*, **38**, 709 ~ 719. [20]

8 **Arnold, S.J.** and **Wade, M.J.** (1984*b*) On the measurement of natural and sexual selection. Applications. *Evolution*, **38**, 720 ~ 734. [20]

9 **Atkins, K.D.** and **Thompson, R.** (1986) Predicted and realized responses to selection for an index of bone length and body weight in Scottish Blackface sheep. *Anim. Prod.*, **43**, 421 ~ 435. [11]

10 **Ayala, F.J.** and **Campbell, C.A.** (1974) Frequency - dependent selection. *Ann. Rev. Ecology and Systematics*, **5**, 115 ~ 138. [2]

11 **Bailey, D.W.** (1959) Rates of snbline divergence in highly inbred strains of mice. *J. Hered.*, **50**, 26 ~ 30. [15]

12 **Baptist, R.** and **Robertson, A.** (1976) Asymmetrical responses to automatic selection for body size in *Drosophila melanogaster*. *Theor. Appl. Genet.*, **47**, 209 ~ 213. [12]

13 **Barker, J. S. F.** (1979) Inter-locus interactions: a review of experimental evidence. *Theor. Pop. Biol.*, **16**, 323 ~ 346. [8]

14 **Barker, J.S.F.** and **Robertson, A.** (1966) Genetic and phenotypic parameters for the first three lactations in Friesian cows. *Anim. Prod.*, **8**, 221 ~ 240. [8, 10, 19]

15 **Bartlett, M.S.** and **Haldane, J.B.S.** (1935) The theory of inbreeding with forced heterozygosis. *J. Genet.*, **31**, 327 ~ 340. [5]

16 **Barton, N.** (1990) Pleiotropic models of quantitative variation. *Genetics*, **124**, 773 ~ 782. [20]

17 **Barton, N.** and **Turelli, M.** (1989) Evolutionary quantitative genetics: How little do we know? *Annu. Rev. Genet.*, **23**, 337 ~ 370. [20]

18 **Becker, W.A.** (1984) *Manual of Quantitative Genetics* (4th edn). Academic Enterprises, Pullman, Wash., USA. [绪论, 13, 附录 B]

19 **Beckmann, J.S.** and **Soller, M.** (1988) Detection of linkage between marker loci and loci affecting quantitative traits in crosses between segregating populations. *Theor. Appl. Genet.*, **76**, 228 ~ 236. [21]

20 **Bereskin, B.** *et al.* (1968) Inbreeding and swine productivity traits. *J. Anim. Sci.*, **27**, 339 ~ 350. [14]

21 **Berger, E.** (1976) Heterosis and the maintenance of enzyme polymorphism. *Am. Nat.*, **110**, 823 ~ 839. [2]

22 **Bohren, B.B., Hill, W.G.** and **Robertson, A.** (1966) Some observations on asymmetrical correlated responses. *Genet. Res.*, **7**, 44 ~ 57. [19]

23 **Bohren, B.B., McKean, H.E.** and **Yamada, Y.** (1961) Relative efficiencies of heritability estimates based on regression of offspring on parent. *Biometrics*, **17**, 481 ~ 491. [10]

24 **Bowman, J.C.** and **Falconer, D.S.** (1960) Inbreeding depression and heterosis of litter size in mice. *Genet. Res.*, **1**, 262 ~ 274. [14]

25 **Bradford, G.E.** and **Famula, T.R.** (1984) Evidence for a major gene for rapid postweaning growth in mice. *Genet. Res.*, **44**, 293 ~ 308. [21]

26 **Breese, E.L.** and **Mather, K.** (1957) The organization of polygenic activity within a chromosome in *Drosophila*. I. Hair characters. *Heredity*, **11**, 373 ~ 395. [21]

27 **Briscoe, D.A., Robertson, A.** and **Malpica, J. - M.** (1975) Dominance at the *Adh* locus in response of adult *Drosophila melanogaster* to environmental alcohol. *Nature*, **255**, 148 ~ 149. [2]

28 **Brumby, P.J.** (1958) Monozygotic twins and dairy cattle improvement. *Anim. Breed. Abstr.*, **26**, 1 ~ 12. [10]

29 **Bulmer, M.G.** (1971) The effect of selection on genetic variability. *Am. Nat.*, **105**, 201 ~ 211. [20]

30 **Bulmer, M.G.** (1976) The effects of selection on genetic variability: a simulation study. *Genet. Res.*, **28**, 101 ~ 117. [20]

31 **Bulmer, M.G.** (1985) *The Mathematical Theory of Quantitative Genetics*. Clarendon Press, Oxford. [绪论, 11, 20]

32 **Bulmer, M.G.** (1989) Maintenance of genetic variability by mutation - selection balance: a child's guide through the jungle. *Genome*, **31**, 761 ~ 767. [20]

33 **Bünger, L.** and **Herrendörfer, G.** (1994) Analysis of a long - term selection experiment with an exponential model. *J. Anim. Breed. Genet.*, **111**, 1 ~ 13. [12]

34 **Bünger, L., Renne, U.** and **Dietl, G.** (1994) 60 generations selection for an index combining high body weight and high stress resistance in laboratory mice. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, Vol.19, 16 ~ 19. [12]

35 **Buri, P.** (1956) Gene frequency in small populations of mutant *Drosophila*. *Evolution*, **10**, 367 ~ 402. [3]

36 **Caballero, A.** (1994) Developments in the prediction of effective population size. *Heredity*, **73**, 657 ~ 679. [4]

37 **Caballero, A.** and **Keightley, P.D.** (1994) A pleiotropic nonadditive model of variation in quantitative traits. *Genetics*, **138**, 883 ~ 900. [20]

38 **Calhoon, R.E.** and **Bohren, B.B.** (1974) Genetic gains from reciprocal recurrent and within ~ line selection for egg production in the fowl. *Theor. Appl. Genet.*, **44**, 364 ~ 372. [16]

39 **Chai, C.K.** (1957) Developmental homeostasis of body growth in mice. *Am. Nat.*, **91**, 49 ~ 55. [15]

40 **Chi, R.K., Eberhart, S.A.** and **Penny, L.H.** (1969) Covariances among relatives in a maize variety (*Zea mays* L.). *Genetics*, **63**, 511 ~ 520. [9]

41 **Christian, J.C., Kang, K.W.** and **Norton, J.A.** (1974) Choice of an estimate of genetic variance from twin data. *Am. J. Hum. Genet.*, **26**, 154 ~ 161. [10]

42 **Clark, A.G., Wang, L.** and **Hulleberg, T.** (1995) Spontaneous mutation rate of modifiers of metabolism in *Drosophila*. *Genetics*, **139**, 767 ~ 779. [20]

43 **Clarke, B.** (1969) The evidence for apostatic selection. *Heredity*, **24**, 347 ~ 352. [2]

44 **Clarke, B.C.** (1975) The contribution of ecological genetics to evolutionary theory: Detecting the direct effects of natural selection on particular polymorphic loci. *Genetics*, **79**, 101 ~ 113. [2]

45 **Clarke, B.C.** (1979) The evolution of genetic diversity. *Proc. R. Soc. Lond. B.*, **205**, 453 ~ 474. [2]

46 **Clayton, G.A., Morris, J.A.** and **Robertson, A.** (1957) An experimental check on quantitative genetical theory. Ⅰ. Short – term responses to selection. *J. Genet.*, **55**, 131 ~ 151. [8, 10, 11, 12, 13]

47 **Clayton, G.A.** and **Robertson, A.** (1955) Mutation and quantitative variation. *Am. Nat.*, **89**, 151 ~ 158. [8, 20]

48 **Clayton, G.A.** and **Robertson, A.** (1957) An experimental check on quantitative genetical theory. Ⅱ. The long – term effects of selection. *J. Genet.*, **55**, 152 ~ 170. [12, 13]

49 **Cloninger, C.R., Rice, J.** and **Reich, T.** (1979) Multifactorial inheritance with cultural transmission and assortative mating. Ⅱ. A general model of combined polygenic and cultural inheritance. *Am. J. Hum. Genet.*, **31**, 176 ~ 198. [10]

50 **Clutton – Brock, T.H., Guinness, F.E.** and **Albon, S.D.** (1982) *Red Deer: Behaviour and Ecology of Two Sexes*. University of Chicago Press, Chicago, Ill., USA. [20]

51 **Cockerham, C.C.** (1954) An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics*, **39**, 859 ~ 882. [8]

52 **Cockerham, C.C.** (1956) Effects of linkage on the covariances between relatives. *Genetics*, **41**, **138 ~ 141.** [9]

53 **Cockerham, C.C.** (1963) Estimation of genetic variances. pp.53 ~ 94 in Hanson, W.D. and Robinson, H.F. (eds) *Statistical Genetics and Plant Breeding*. Nat. Acad. Sci. Nat. Res. Council Publ. No.982, Washington, D.C., USA. [8]

54 **Comstock, R.E., Robinson, H.F.** and **Harvey, P.H.** (1949) A breeding procedure designed to make maximum use of both general and specific combining ability. *J. Amer. Soc. Agron.*, **41**, 360 ~ 367. [16]

55 **Corder, E.H., Saunders, A.M., Strittmatter, W.J., Schmechel, D.E., Gaskell, P.C., Small, G.W., Roses, A.D., Haines, J.L.** and **Pericak – Vance, M.A.** (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*, **261**, 921 ~ 923. [21]

56 **Cornelius, P.L.** and **Dudley, J.W.** (1974) Effects of inbreeding by selfing and full – sib mating in a maize population. *Crop Sci.*, **14**, 815 ~ 819. [14]

57 **Crow, J.F.** (1948) Alternative hypotheses of hybrid vigor. *Genetics*, **33**, 477 ~ 487. [16]

58 **Crow, J.F.** (1952) Dominance and overdominance. pp.282 ~ 297 in Gowen, J.W. (ed.), *Heterosis*, Iowa State College, Ames, Iowa, USA. [16, 20]

59 **Crow, J.F.** (1957) Possible consequences of an increased mutation rate. *Eugen. Quart.*, **4**, 67 ~ 80. [2]

60 **Crow, J.F.** (1986) *Basic Concepts in Population, Quantitative, and Evolutionary Genetics*. Freeman, New York, USA. [绪论, 1, 2]

61 **Crow, J.F.** (1989) Fitness variation in natural populations. pp.91 ~ 97 in Hill, W.G. and Mackay, T.F.C. (eds), *Evolution and Animal Breeding*. C.A.B. International, Wallingford, UK. [20]

62 **Crow, J.F.** and **Kimura, M.** (1970) *An Introduction to Population Genetics Theory*, Harper and Row, New York. [绪论, 1, 4, 5, 8, 9, 10, 14, 15]

63 **Crow, J.F.** and **Nagylaki, T.** (1976) The rate of change of a character correlated with fitness. *Amer. Nat.*, **110**, 207 ~ 213. [20]

64 **Crow, J.F.** and **Simmons, M.J.** (1983) The mutation load in *Drosophila*. pp.1 ~ 35 in Carson, H.L., Ashburner, M. and Thompson, J.N. (eds), *The Genetics and Biology of Drosophila*. Academic Press, New York, NY, USA. [20]

65 **Curnow, R.N.** (1964) The effect of continued selection of phenotypic intermediates on gene frequency. *Genet. Res.*, **5**, 341 ~ 353. [20]

66 **Curnow, R.N.** (1984) Progeny testing for all – or – none traits when a multi – factorial model applies. *Biometrics*, **40**, 375 ~ 382. [18]

67 **Curnow, R.N.** and **Smith, C.** (1975) Multifactorial models for familial diseases in man. *J. Roy. Statist. Soc. A*, **138**, 131 ~ 169. [18]

68 **Davies, J.L., Kawaguchi, Y., Bennett, S.T., Copeman, J.B., Cordell, H.J., Pritchard, L.E., Reed, P.W., Gough, S.C.L., Jenkins, S.C., Palmer, S.M., Bal-**

four, **K.M.**, **Rowe**, **B.R.**, **Farrall**, **M.**, **Barnett**, **A.H.**, **Bain**, **S.C.** and **Todd**, **J.A.** (1994) A genome – wide search for human type 1 diabetes susceptibility genes. *Nature*, **371**, 130 ~ 136. [21]

69 **Davisson**, **M.T.** and **Roderick**, **T.H.** (1989) Linkage Map. pp.416 ~ 427 in Lyon, M.F. and Searle, A.G. (eds), *Genetic Variants and Strains of the Laboratory Mouse*, 2nd edn. Oxford University Press, Oxford, Fisher Verlag, Stuttgart. [5]

70 **Dean**, **A.M.**, **Dykhuizen**, **D.E.** and **Hartl**, **D.L.** (1988) Theories of metabolic control in quantitative genetics. Ch.49 in Weir, B.S., Eisen, E.J., Goodman, M.M. and Namkoong, G. (eds), *Proc. 2nd Internat. Conf. on Quantitative Genetics*. Sinauer, Sunderland, Mass., USA. [21]

71 **Dempster**, **E.R.** and **Lerner**, **I.M.** (1950) Heritability of threshold characters. *Genetics*, **35**, 212 ~ 236. [18]

72 **deVicente**, **M.C.** and **Tanksley**, **S.D.** (1993) QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics*, **134**, 585 ~ 596. [21]

73 **Dickerson**, **G.E.** (1952) Inbred lines for heterosis tests? pp.330 ~ 351 in Gowen, J.W. (ed.), *Heterosis*, Iowa State College, Ames, Iowa, USA. [16]

74 **Dickerson**, **G.E.** (1962) Implications of genetic – environmental interaction in animal breeding. *Anim. Prod.*, **4**, 47 ~ 63. [19]

75 **Dickerson**, **G.E.** (1969) Experimental approaches in utilising breed resources. *Anim. Breed. Abstr.*, **37**, 191 ~ 202. [16]

76 **Dobzhansky**, **Th.** and **Pavlovsky**, **O.** (1955) An extreme case of heterosis in a Central American population of *Drosophila tropicalis*. *Proc. Natnl. Acad. Sci. Wash.*, **41**, 289 ~ 295. [2]

77 **Doebley**, **J.** and **Stec**, **A.** (1991) Genetic analysis of the morphological differences between maize and teosinte. *Genetics*, **129**, 285 ~ 295. [21]

78 **Doebley**, **J.** and **Stec**, **A.** (1993) Inheritance of the morphological differences between maize and teosinte: Comparison of results for two F_2 populations. *Genetics*, **134**, 559 ~ 570. [21]

79 **Dudley**, **J.W.** (1977) 76 generations of selection for oil and protein percentage in maize. pp.459 ~ 473 in Pollak, E., Kempthorne, O. and Bailey, T.B. (eds), *Proc. Int. Conf. Quantitative Genetics*, Iowa State Univ., Ames, Iowa, USA. [12]

80 **Dudley**, **J.W.** and **Lambert**, **J.R.** (1992) Ninety generations of selection for oil and protein content in maize. *Maydica*, **37**, 1 ~ 7. [12]

81 **Eaves**, **L.** (1976) The effect of cultural transmission on continuous variation. *Heredity*, **37**, 41 ~ 57. [10]

82 **Eaves**, **L.J.**, **Last**, **K.A.**, **Young**, **P.A.** and **Martin**, **N.G.** (1978) Model – fitting approaches to the analysis of human behaviour. *Heredity*, **41**, 249 ~ 320. [10]

83 **Eberhart**, **S.A.** (1977) Quantitative genetics and practical corn breeding. pp.491 ~ 502 in Pollak, E., Kempthorne, O. and Bailey, T.B. (eds), *Proc. Int. Conf. Quantitative Genetics*, Iowa State Univ., Ames, Iowa, USA. [16]

84 **Edwards, A.W.F.** (1994) The fundamealnt theorem of natural selection. *Biol. Rev.*, **69**, 443 ~ 474. [20]

85 **Edwards, M. D., Stuber, C. W.** and **Wendel, J. F.** (1987) Molecular – marker – facilitated investigations of quantitative trait loci in maize. I. Numbers, genomic distburition and types of gene action. *Genetics*, **116**, 113 ~ 125. [21]

86 **Eisen, E. J.** (1980) Conclusions from long – term selection experiments with mice. *Z. Tierzüchtg. Züchtgsbiol.*, **97**, 305 ~ 319. [12]

87 **Eklund, J.** and **Bradford, G.E.** (1977) Genetic analysis of a strain of mice plateaued for litter size. *Genetics*, **85**, 529 ~ 542. [12, 14]

88 **Elston, R. C.** (1988) Human quantitative genetics. Ch.25 in Weir, B.S., Eisen, E.J., Goodman, M.M. and Namkoong, G. (eds), *Proc. 2nd Internat. Conf. on Quantitative Genetics*. Sinauer, Sunderland, Mass., USA. [10]

89 **Emery, A.E.H.** (1976) *Methodology in Medical Genetics*. Churchill Livingstone, Edinburgh. [18]

90 **Emigh, T.H.** and **Pollak, E.** (1979) Fixation probabilities and effective population numbers in diploid populations with overlapping generations. *Theor. Pop. Biol.*, **15**, 86 ~ 107. [4]

91 **Emsley, A., Dickerson, G.E.** and **Kashyap, T.S.** (1977) Genetic parameters in progeny – test selection for field performance of strain – cross layers. *Poult. Sci.*, **56**, 121 ~ 146. [10, 19]

92 **Endler, J.A.** (1986) *Natural Selection in the Wild*. Princeton University Press, Princeton, NJ, USA. [2, 20]

93 **Enfield, F.D.** (1977) Selection experiments in tribolium designed to look at gene action issues. pp.177 ~ 190 in Pollak, E., Kempthorne, O. and Bailey, T.B. (eds), *Proc. Int. Conf. Quantitative Genetics*, Iowa State Univ., Ames, Iowa, USA. [12]

94 **England, F.** (1977) Response to family selection based on replicated trials. *J. Agric. Sci.*, **88**, 127 ~ 134. [13]

95 **Falconer, D.S.** (1953) Selection for large and small size in mice. *J. Genet.*, **51**, 470 ~ 501. [11, 20]

96 **Falconer, D.S.** (1954) Asymmetrical responses in selection experiments. pp.16 ~ 41 in *Symposium on Genetics of Population Structure*. Internat. Union Biol. Sci., Naples, Series B, No.15. [2, 11, 12, 17]

97 **Falconer, D.S.** (1955) Patterns of response in selection experiments with mice. *Cold Spring Harbor Symp. Quant. Biol.*, **20**, 178 ~ 196. [11, 12]

98 **Falconer, D.S.** (1960*a*) The genetics of litter size in mice. *J. Cell. Comp. Physiol.*, **56** (*Suppl.*1), 153 ~ 167. [11, 14]

99 **Falconer, D.S.** (1960*b*) Selection of mice for growth on high and low planes of nutrition. *Genet. Res.*, **1**, 91 ~ 113. [19]

100 **Falconer, D.S.** (1963) Quantitative inheritance. pp.193 ~ 216 in Burdette, W.J. (ed.), *Methodology in Mammalian Genetics*, Holden – Day, San Francisco. [10]

101 **Falconer, D.S.** (1965*a*) The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann. Hum. Genet.*, **29**, 51 ~ 76. [18, 附录 A]

102 **Falconer, D.S.** (1965*b*) Maternal effects and selection response. pp. 763 ~ 774 in Geerts, S.J. (ed.), *Genetics Today*, Proc. XI Internat. Congr. Genetics, Vol. 3, Pergamon, Oxford. [9, 10, 11, 12, 20]

103 **Falconer, D.S.** (1967) The inheritance of liability to diseases with variable age of onset, with particular reference to diabetes mellitus. *Ann. Hum. Genet.*, **31**, 1 ~ 20. [18]

104 **Falconer, D.S.** (1971) Improvement of litter size in a strain of mice at a selection limit. *Genet. Res.*, **17**, 215 ~ 235. [5, 12, 14]

105 **Falconer, D.S.** (1973) Replicated selection for body weight in mice. *Genet. Res.*, **22**, 291 ~ 321. [4, 10, 12]

106 **Falconer, D.S.** (1977) Some results of the Edinburgh selection experiments with mice. pp. 101 ~ 115 in Pollak, E., Kempthorne, O. and Bailey, T.B. (eds), *Proc. Int. Conf. Quantitative Genetics*, Iowa State Univ., Ames, Iowa, USA. [12]

107 **Falconer, D.S.** (1985) A note on Fisher's 'average effect' and 'average excess'. *Genet. Res.*, **46**, 337 ~ 347. [7]

108 **Falconer, D.S.** (1990) Selection in different environments: effects on environmental sensitivity (reaction norm) and on mean performance. *Genet. Res.*, **56**, 57 ~ 70. [19]

109 **Famula, T.R.** (1986) Identifying single genes of large effect in quantitative traits using best linear unbiased prediction. *J. Anim. Sci.*, **63**, 68 ~ 76. [21]

110 **Fatokun, C.A., Menancio – Hautea, D.I., Danesh, D.** and **Young, N.D.** (1992) Evidence for orthologous seed weight genes in cowpea and mung bean based on RFLP mapping. *Genetics*, **132**, 841 ~ 846. [21]

111 **Felsenstein, J.** (1976) The theoretical population genetics of variable selection and migration. *Ann. Rev. Genet.*, **10**, 253 ~ 280. [2]

112 **Festing, M.** (1973) A multivariate analysis of snhline divergence in the shape of the mandible in C57BL/Gr mice. *Genet. Res.*, **21**, 121 ~ 132. [15]

113 **Fisher, R.A.** (1918) The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinburgh*, **52**, 399 ~ 433. [绪论, 7]

114 **Fisher, R.A.** (1941) Average excess and average effect of a gene substitution. *Ann. Eugen. (Lond.)*, **11**, 53 ~ 63. [7]

115 **Fisher, R.A.** and **Yates, F.** (1963) *Statistical Tables* (6th edn). Longman, London and New York. [11]

116 **Fowler, V.R., Bichard, M.** and **Pease, A.** (1976) Objectives in pig breeding. *Anim. Prod.*, **23**, 365 ~ 387. [19]

117 **Frank, S.A.** and **Slatkin, M.** (1992) Fisher's fundamental theorem of natural selection. *Trends Ecol. Evol.*, **7**, 92 ~ 95. [20]

118 **Frankham, R.** (1990) Are responses to artificial selection for reproductive fitness characters consistently asymmetrical? *Genet. Res.*, **56**, 35 ~ 42. [12]

119 **Frankham, R.** (1995) Effcctivc population size/adult population size ratios in wildlife: a review. *Genet. Res.*, **66**, in press. [4]

120 **Fredeen, H.T.** and **Jonsson, P.** (1957) Genic variance and covariance in Danish Landrace swine as evaluated under a system of individual feeding of progeny test groups. *Zeits. Tierzüch. u Züicht. Biol.*, **70**, 348 ~ 363. [10]

121 **Gama, E.E.G.** and **Hallauer, A.R.** (1977) Relation between inbred and hybrid traits in maize. *Crop Sci.*, **17**, 703 ~ 706. [16]

122 **Gardner, C.O.** (1977) Quantitative genetic studies and population improvement in maize and sorghum. pp. 475 ~ 489 in Pollak, E., Kempthorne, O. and Bailey, T.B. (eds), *Proc. Int. Conf. Quantitative Genetics*, Iowa State Univ., Ames, Iowa, USA. [16]

123 **Garnett, I.** and **Falconer, D.S.** (1975) Protein variation in strains of mice differing in body size. *Genet. Res.*, **25**, 45 ~ 47. [4]

124 **Geiger, H.H.** (1988) Epistasis and heterosis. Ch. 34 in Weir, B.S., Eisen, E.J., Goodman, M.M. and Namkoong, G. (eds), *Proc. 2nd Internat. Conf. on Quantitative Genetics*. Sinauer, Sunderland, Mass., USA. [14, 16]

125 **Gibson, J.B.** and **Bradley, B.P.** (1974) Stabilizing selection in constant and fluctuating environments. *Heredity*, **33**, 293 ~ 302. [20]

126 **Gimelfarb, A.** (1986) Offspring – parent genotypic regression: how linear is it? . *Biometrics*, **42**, 67 ~ 71. [11]

127 **Gowe, R.S., Robertson, A.** and **Latter, B.D.H.** (1959) Environment and poultry breeding problems. 5. The design of poultry control strains. *Poult. Sci.*, **38**, 462 ~ 471. [4]

128 **Grant, P.R.** (1986) *Ecology and Evolution of Darwin's Finches*. Princeton University Press, Princeton, NJ, USA. [20]

129 **Greaves, J.H., Redfern, R., Ayres, P.B.** and **Gill, J.E.** (1977) Warfarin resistance: a balanced polymorphism in the Norway rat. *Genet. Res.*, **30**, 257 ~ 263. [2]

130 **Green, E.L.** (1962) Quantitative genetics of skeletal variations in the mouse. Ⅱ. Crosses between four inbred strains. *Genetics*, **47**, 1085 ~ 1096. [18]

131 **Green, E.L.** (ed.) (1968) *Handbook of Genetically Standardized JAX Mice*. The Jackson Laboratory, Bar Harbor, Maine. [14]

132 **Grewal, M.J.** (1962) The rate of genetic divergence of sublines in the C57BL strain of mice. *Genet. Res.*, **3**, 226 ~ 237. [15]

133 **Griffing, B.** (1956*a*) A generalised treatment of the use of diallel crosses in quantitative inheritance. *Heredity*, **10**, 31 ~ 50. [15]

134 **Griffing, B.** (1956*b*) Concept of general and specific combining ability in relation to diallel crossing systems. *Aust. J. Biol. Sci.*, **9**, 463 ~ 493. [15]

135 **Grundbacher, F.J.** (1974) Heritability estimates and genetic and environmental correlations for the human immunoglobulins G, M, and A. *Am. J. Hum. Genet.*, **26**, 1 ~ 12. [10, 19]

136 **Haldane, J.B.S.** (1932) *The Causes of Evolution*. Longmans, Green, London. [绪论]

137 **Haldane, J.B.S.** (1936) The amount of heterozygosis to be expected in an approximately

pure line. *J. Genet.*, **32**, 375 ~ 391. [5]

138 **Haldane, J.B.S.** (1949) The rate of mutation of human genes. Proc. 8th Internat. Congr. Genetics, *Hereditas*, *Suppl. Vol.* 1949, 267 ~ 273. [2]

139 **Haldane, J.B.S.** (1955) The complete matrices for brother – sister and alternate parent – offspring mating involving one locus. *J. Genet.*, **53**, 315 ~ 324. [5]

140 **Haley, C.S.** and **Knott, S.A.** (1992) A simple regression model for interval mapping in line crosses. *Heredity*, **69**, 315 ~ 324. [21]

141 **Hallauer, A.R.** and **Miranda, J.B.** (1981) *Quantitative Genetics in Maize Breeding*, Iowa State Univ., Ames, Iowa, USA. [6]

142 **Hallauer, A.R.** and **Sears, J.H.** (1973) Changes in quantitative traits associated with inbreeding in a synthetic variety of maize. *Crop Sci.*, **13**, 327 ~ 330. [14]

143 **Halliburton, R.** and **Gall, G.A.E.** (1981) Disruptive selection and assortative mating in *Tribolium castaneum*. *Evolution*, **35**, 829 ~ 843. [20]

144 **Hammond, K.** and **James, J.W.** (1970) Genes of large effect and the shape of the distribution of a quantitative character. *Aust. J. Biol. Sci.*, **23**, 867 ~ 876. [21]

145 **Hammond, K.** and **Nicholas, F.W.** (1972) The sampling variance of the correlation coefficients estimated from two – fold nested and offspring – parent regression analyses. *Theor. Appl. Genet.*, **42**, 97 ~ 100. [19]

146 **Hancock, J.** (1954) Monozygotic twins in cattle. *Adv. Genet.*, **6**, 141 ~ 181. [10]

147 **Hanset, R.** (1982) Major genes in animal production, examples and perspectives: Cattle and pigs. *Proceedings of the 2nd World Congress on Genetics Applied to Livestock Production*, Vol. 6, 439 ~ 453. [21]

148 **Hanset, R.** and **Michaux, C.** (1985*a*) On the genetic determinism of muscular hypertrophy in the Belgian White and Blue cattle breed. I. *Génét. Sél. Evol.*, **17**, 359 ~ 368. [21]

149 **Hanset, R.** and **Michaux, C.** (1985*b*) On the genetic determinism of muscular hypertrophy in the Belgian White and Blue cattle breed. II. *Génét. Sél. Evol.*, **17**, 369 ~ 386. [21]

150 **Hardy, G.H.** (1908) Mendelian proportions in a mixed population. *Science*, **28**, 49 ~ 50. [1]

151 **Hartl, D.L.** and **Clark, A.G.** (1989) *Principles of Population Genetics*. Sinauer, Sunderland, Mass., USA. [绪论, 4]

152 **Hartl, D.L.**, **Dykhuizen, D.E.** and **Dean, A.M.** (1985) Limits of adaptation in the evolution of selective neutrality. *Genetics*, **111**, 635 ~ 674. [21]

153 **Hayman, B.I.** and **Mather, L.** (1953) The progress of inbreeding when homozygotes are at a disadvantage. *Heredity*, **7**, 165 ~ 183. [5]

154 **Hedrick, P.W.** (1985) *Genetics of Populations*. Jones and Barlett, Boston, USA. [绪论]

155 **Henderson, C.R.** (1977) Prediction of future records. pp. 615 ~ 638 in Pollak, E., Kempthorne, O. and Bailey, T.B. (eds), *Proc. Int. Conf. Quantitative Genetics*. Iowa State Univ., Ames, Iowa, USA. [13]

156 **Henderson, C.R.** (1988) Progress in statistical methods applied to quantitative genetics since

1976. Ch. 8 in Weir, B.S., Eisen, E.J., Goodman, M.M. and Namkoong, G. (eds), *Proc. 2nd Internat. Conf. on Quantitative Genetics*. Sinauer, Sunderland, Mass., USA. [13]

157 **Hill, W.G.** (1971) Design and efficiency of selection experiments for estimating genetic parameters. *Biometrics*, **27**, 293 ~ 311. [12]

158 **Hill, W.G.** (1972*a*) Estimation of genetic change. Ⅰ. General theory and design of control populations. *Anim. Breed. Abstr.*, **40**, 1 ~ 15. [11]

159 **Hill, W.G.** (1972*b*) Estimation of genetic change. Ⅱ. Experimental evaluation of control populations. *Anim. Breed. Abstr.*, **40**, 193 ~ 213. [11]

160 **Hill, W.G.** (1972*c*) Estimation of realized heritabilities from selection experiments. Ⅰ. Divergent selection. *Biometrics*, **28**, 747 ~ 765. [12]

161 **Hill, W.G.** (1972*d*) Estimation of realized heritabilities from selection experiments. Ⅱ. Selection in one direction. *Biometrics*, **28**, 767 ~ 780. [12]

162 **Hill, W.G.** (1979) A note on effective population size with overlapping generations. *Genetics*, **92**, 317 ~ 322. [4]

163 **Hill, W.G.** (1980) Design of quantitative genetic selection experiments. pp. 1 ~ 13 in Robertson, A. (ed.), *Selection Experiments in Laboratory and Domestic Animals*. Commonwealth Agricultural Bureaux, Slough, UK. [12]

164 **Hill, W.G.** (1982*a*) Rates of change in quantitative traits from fixation of new mutations. *Proc. Natnl. Acad. Sci., Wash.*, **79**, 142 ~ 145. [12]

165 **Hill, W.G.** (1982*b*) Predictions of response to artificial selection from new mutations. *Genet. Res.*, **40**, 255 ~ 278. [12, 20]

166 **Hill, W.G.** (1984) *Quantitative Genetics*, Pts Ⅰ and Ⅱ. Benchmark Papers in Genetics, Vol. 15. Van Nostrand Reinhold, New York, USA. [绪论]

167 **Hill, W.G.**, **Caballero, A.** and **Keightley, P.D.** (1994) Variation from spontaneous mutation for body size in the mouse. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, Vol. 19, 67 ~ 70. [20]

168 **Hill, W.G.** and **Keightley, P.D.** (1988) Interrelations of mutation, population size, artificial and natural selection. Ch. 6 in Weir, B.S., Eisen, E.J., Goodman, M.M. and Namkoong, G. (eds), *Proc. 2nd Internat. Conf. on Quantitative Genetics*. Sinauer, Sunderland, Mass., USA. [12]

169 **Hill, W.G.** and **Knott, S.** (1990) Identification of genes with large effects. pp. 477 ~ 494 in Gianola, D. and Hammond, K. (eds), *Advances in Statistical Methods for Genetic Improvement of Livestock*. Springer – Verlag, Berlin, Germany. [21]

170 **Hill, W.G.** and **Nicholas, F.W.** (1974) Estimation of heritability by both regression of offspring on parent and intra – class correlation of sibs in one experiment. *Biometrics*, **30**, 447 ~ 468. [10]

171 **Hill, W.G.** and **Rasbash, J.** (1986) Models of long term artificial selection in finite population. *Genet. Res.*, **48**, 41 ~ 50. [12]

172 **Hill, W.G.** and **Robertson, A.** (1966) The effect of linkage on limits to artificial selection. *Genet. Res.*, **8**, 269 ~ 294. [12]

173 **Hill, W.G.** and **Robertson, A.** (1968) The effects of inbreeding at loci with heterozygote advantages. *Genetics*, **60**, 615 ~ 628. [14]

174 **Hill, W.G.** and **Thompson, R.** (1977) Design of experiments to estimate offspring – parent regression using selected parents. *Anim. Prod.*, **24**, 163 ~ 168. [10]

175 **Hoekstra, R.F.**, **Bijlsma, R.** and **Dolman, A.J.** (1985) Polymorphism from environmental heterogeneity: models are only robust if the heterozygote is close in fitness to the favoured homozygote in each environment. *Genet. Res.*, **45**, 299 ~ 314. [2]

176 **Hollingsworth, M.J.** and **Maynard Smith, J.** (1955) The effects of inbreeding on rate of development and on fertility in *Drosophila subobscura*. *J. Genet.*, **53**, 295 ~ 314. [14]

177 **Houle, D.** (1992) Comparing evolvability and variability of quantitative traits. *Genetics*, **130**, 195 ~ 204. [20]

178 **Houle, D.**, **Hughes, K.A.**, **Hoffmaster, D.K.**, **Ihara, J.**, **Assimacopoulos, S.**, **Canada, D.** and **Charlesworth, B.** (1994) The effects of spontaneous mutation on quantitative traits. I. Variance and covariance of life – history traits. *Genetics*, **138**, 773 ~ 785. [20]

179 **Hudson, R.R.**, **Kreitman, M.** and **Aguadé, M.** (1987) A test of neutral molecular evolution based on nucleotide data. *Genetics*, **116**, 153 ~ 159. [2]

180 **Hunt, H.R.**, **Hoppert, C.A.** and **Erwin, W.G.** (1944) Inheritance of susceptibility to caries in albino rats (*Mus norvegicus*). *J. Dental Res.*, **23**, 385 ~ 401. [17]

181 **Huntley, R.M.C.** (1966) Heritability of intelligence. pp. 201 ~ 218 in Meede, J.E. and Parkes, A.S. (eds), *Genetic and Environmental Factors in Human Ability*, Oliver and Boyd, Edinburgh. [10]

182 **Hyde, J.S.** (1973) Genetic homeostasis and behaviour: analysis, data and theory. *Behaviour Genetics*, **3**, 233 ~ 245. [15]

183 **Jacquard, A.** (1974) *The Genetic Structure of Populations* (translated by D. and B. Charlesworth). Springer – Verlag, Berlin. [绪论]

184 **James, J.W.** (1961) Selection in two environments. *Heredity*, **16**, 145 ~ 152. [19]

185 **Jansen, R.C.** and **Stam, P.** (1994) High resolution mapping of quantitative traits into multiple loci via interval mapping. *Genetics*, **136**, 1447 ~ 1455. [21]

186 **Jinks, J.L.** and **Pooni, H.S.** (1976) Predicting the properties of recombinant inbred lines derived by single seed descent. *Heredity*, **36**, 253 ~ 266. [16]

187 **Jódar, B.** and **López – Fanjul, C.** (1977) Optimum proportions selected with unequal sex numbers. *Theor. Appl. Genet.*, **50**, 57 ~ 61. [12]

188 **Johnson, L.L.** (1981) At how many histocompatibility loci do congenic mouse strains differ? Probability estimates and some implications. *J. Hered.*, **72**, 27 ~ 31. [5]

189 **Jones, L.P.**, **Frankham, R.** and **Barker, J.S.F.** (1968) The effects of population size and selection intensity in selection for a quantitative character in *Drosophila*. Ⅱ. Long – term response to selection. *Genet. Res.*, **12**, 249 ~ 266. [12]

190 **Kacser, H.** and **Burns, J.A.** (1981) The molecular basis of dominance. *Genetics*, **97**, 639 ~ 674. [21]

191 **Kammerer, C.M.**, **MacCluer, J.W.** and **Bridges, J.M.** (1984) An evaluation of three statistics of structured exploratory data analysis. *Am. J. Hum. Genet.*, **36**, 187 ~ 196. [21]

192 **Karlin, S.**, **Carmelli, D.** and **Williams, R.** (1979) Index measures for assessing the mode of inheritance of continuously distributed traits. I. Theory and justifications. *Theor. Pop. Biol.*, **16**, 81 ~ 106. [21]

193 **Karlin, S.**, **Williams, P.T.** and **Carmelli, D.** (1981) Structured exploratory data analysis (SEDA) for determining mode of inheritance of quantitative traits. I. Simulation studies on the effect of background distributions. *Am. J. Hum. Genet.*, **33**, 262 ~ 281. [21]

194 **Kaufman, P.K.**, **Enfield, F.D.** and **Comstock, R.E.** (1977) Stabilizing selection for pupa weight in *Tribolium castaneum*. *Genetics*, **87**, 327 ~ 341. [20]

195 **Kearsey, M.J.** and **Barnes, B.W.** (1970) Variation for metrical characters in *Drosophila* populations. Ⅱ. Natural selection. *Heredity*, **25**, 11 ~ 21. [20]

196 **Kearsey, M.J.** and **Kojima, K.** (1967) The genetic architecture of body weight and egg hatchability in *Drosophila melanogaster*. *Genetics*, **56**, 23 ~ 37. [8, 20]

197 **Keightley, P. D.** and **Hill, W. G.** (1990) Variation maintained in quantitative traits with mutation-selection balance: Pleiotropic side-effects on fitness traits. *Proc. Roy. Soc. Lond. B*, **242**, 95 ~ 100. [20]

198 **Keightley, P.D.**, **Mackay, T.F.C.** and **Caballero, A.** (1993) Accounting for bias in estimates of the rate of polygenic mutation. *Proc. Roy. Soc. Lond. B*, **253**, 291 ~ 296. [20]

199 **Kempthorne, O.** (1955*a*) The theoretical values of correlations between relatives in random mating populations. *Genetics*, **40**, 153 ~ 167. [8, 9]

200 **Kempthorne, O.** (1955*b*) The correlations between relatives in random mating populations. *Cold Spring Harbor Symp. Quant. Biol.*, **20**, 60 ~ 75. [9]

201 **Kempthorne, O.** (1957) *An Introduction to Genetic Statistics*. Wilcy, New York. [绪论, 8]

202 **Kempthorne, O.** and **Tandon, O.B.** (1953) The estimation of heritability by regression of offspring on parent. *Biometrics*, **9**, 90 ~ 100. [10]

203 **Kennedy, B.W.** (1981) Variance component estimation and prediction of breeding values. *Can. J. Genet. Cytol.*, **23**, 565 ~ 578. [10, 13]

204 **Kennedy, B.W.** and **Sorensen, D.A.** (1988) Properties of mixed model methods for prediction of genetic merit under different genetic models in selected and unselected populations. pp.91 ~ 103 in Weir, B.S., Eisen, E.J., Goodman, M.M. and Namkoong, G. (eds), *Proceedings of the 2nd International Conference on Quantitative Genetics*. Sinauer, Sunderland, Mass., USA. [13]

205 **Kidwell, J.F.** and **Kidwell, M.M.** (1966) The effects of inbreeding on body weight and abdominal chaeta number in *Drosophila melanogaster*. *Can. J. Genet. Cytol.*, **8**, 207 ~ 215. [14]

206 **Kimura, M.** (1955) Solution of a process of random genetic drift with a continuous model. *Proc. Natnl. Acad. Sci. Wash.*, **41**, 144 ~ 150. [3]

207 **Kimura, M.** (1983) *The Neutral Theory of Molecular Evolution*. Cambridge Univ. Press, Cambridge, UK. [绪论, 2, 4, 20]

208 **Kimura, M.** and **Crow, J.M.** (1978) Effect of overall phenotypic selection on genetic change at individual loci. *Proc. Natnl. Acad. Sci., Wash.*, **75**, 6168 ~ 6171. [11]

209 **King, J.W.B.** (1950) Pygmy, a dwarfing gene in the house mouse. *J. Hered.*, **41**, 249 ~ 252. [7]

210 **King, J.W.B.** (1955) Observations on the mutant 'pygmy' in the house mouse. *J. Genet.*, **53**, 487 ~ 497. [7, 17]

211 **Kirkpatrick, M.** and **Lande, R.** (1989) The evolution of maternal effects. *Evolution*, **43**, 485 ~ 503. [11]

212 **Knott, S.A.** and **Haley, C.S.** (1992) Aspects of maximum likelihood methods for the mapping of quantitative trait loci in line crosses. *Genet. Res.*, **60**, 139 ~ 151. [21]

213 **Kondrashov, A.S.** and **Turelli, M.** (1992) Deleterious mutations, apparent stabilizing selection and the maintenance of quantitative variation. *Genetics*, **132**, 603 ~ 618. [20]

214 **Krosigk, C.M. von, Havenstein, G.B., Flock, D.K.** and **McClary, C.F.** (1973) Estimates of response to selection in populations of white leghorns under reciprocal recurrent selection. pp. 256 ~ 271 in *4th Europ. Poult. Conf., London*, 1972, British Poultry Science Ltd. [16]

215 **Lack, D.** (1966) *Population Studies of Birds*. Clarendon Press, Oxford. [20]

216 **Lai, C., Lyman, R.F., Long, A.D., Langley, C.H.** and **Mackay, T.F.C.** (1994) Naturally occurring variation in bristle number associated with DNA sequence polymorphisms at the *scabrous* locus in *Drosophila melanogaster*. *Science*, **266**, 1697 ~ 1702. [1, 21]

217 **Lalouel, J.M., Morton, N.E., MacLean, C.J.** and **Jackson, J.** (1977) Recurrence risks in complex inheritance with special regard to pyloric stenosis. *J. Med. Genet.*, **14**, 408 ~ 414. [18]

218 **Lande, R.** (1975) The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.*, **26**, 221 ~ 235. [20]

219 **Lande, R.** (1981) The minimum number of genes contributing to quantitative variation between and within populations. *Genetics*, **99**, 541 ~ 553. [12]

220 **Lande, R.** and **Arnold, S.J.** (1983) The measurement of selection on correlated characters. *Evolution*, **37**, 1210 ~ 1226. [20]

221 **Lande, R.** and **Thompson, R.** (1990) Efficiency of marker - assisted selection in the improvement of quantitative traits. *Genetics*, **124**, 743 ~ 756. [21]

222 **Lander, E.S.** and **Botstein, D.** (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, **121**, 185 ~ 199. [21]

223 **Latter, B.D.H.** (1965) The response to artificial selection due to autosomal genes of large effect. I. Changes in gene frequency at an additive locus. *Aust. J. Biol. Sci.*, **18**, 585 ~ 598.

[11]

224 **Latter, B.D.H.** (1981) The distribution of heterozygosity in temperate and tropical species of *Drosophila*. *Genet. Res.*, **38**, 137 ~ 156. [2]

225 **Latter, B.D.H.** and **Robertson, A.** (1960) Experimental design in the estimation of heritability by regression methods. *Biometrics*, **16**, 348 ~ 353. [10]

226 **Latter, B.D.H.** and **Robertson, A.** (1962) The effects of inbreeding and artificial selection on reproductive fitness. *Genet. Res.*, **3**, 110 ~ 138. [20]

227 **Lerner, I.M.** (1954) *Genetic Homeostasis*. Oliver and Boyd, Edinburgh. [15, 20]

228 **Lewontin, R.C.** (1964) The interaction of selection and linkage. Ⅱ. Optimum models. *Genetics*, **50**, 757 ~ 782. [20]

229 **Lewontin, R.C.** (1974) *The Genetic Basis of Evolutionary Change*. Columbia Univ., New York. [2, 4]

230 **Li, C.C.** (1976) *First Course in Population Genetics*. Boxwood, Pacific Grove, Calif. USA. [绪论, 5]

231 **Lin, C.Y.** (1978) Index selection for genetic improvement of quantitative characters. *Theor. Appl. Genet.*, **52**, 49 ~ 56. [13]

232 **Lindsley, D.L.** and **Zimm, G.G.** (1992) *The Genome of Drosophila melanogaster*. Academic Press, San Diego, Calif., USA. [21]

233 **Linney, R.**, **Barnes, B.W.** and **Kearsey, M.J.** (1971) Variation for metrical characters in *Drosophila* populations. III. The nature of selection. *Heredity*, **27**, 163 ~ 174. [20]

234 **Livesay, E.A.** (1930) An experimental study of hybrid vigor or heterosis in rats. *Genetics*, **15**, 17 ~ 54. [15]

235 **Long, A.D.**, **Mullaney, S.L.**, **Reid, L.A.**, **Fry, J.D.**, **Langley, C.H.** and **Mackay, T.F.C.** (1995) High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster*. *Genetics*, **139**, 1273 ~ 1291. [6, 21]

236 **Lush, J.L.** (1940) Intra – sire correlations or regressions of offspring on dam as a method of estimating heritability of characteristics. *Am. Soc. Anim. Prod. Proc.*, **33**, 293 ~ 301. [10]

237 **Lush, J.L.** (1947) Family merit and individual merit as bases for selection. *Am. Nat.*, **81**, 241 ~ 261; 362 ~ 379. [13]

238 **Lynch, M.** (1988) The rate of polygenic mutation. *Genet. Res.*, **51**, 137 ~ 148. [12, 15, 20]

239 **Lynch, M.** (1991) The genetic interpretation of inbreeding depression and outbreeding depression. *Evolution*, **45**, 622 ~ 629. [14]

240 **Lynch, M.** and **Hill, W.G.** (1986) Phenotypic evolution by neutral mutation. *Evolution*, **40**, 915 ~ 935. [15, 20]

241 **MacArthur, J.W.** (1949) Selection for small and large body size in the house mouse. *Genetics*, **34**, 194 ~ 209. [12, 17, 20]

242 **Mackay, T.F.C.** (1985*a*) A quantitative genetic analysis of fitness and its components in *Drosophila melanogaster*. *Genet. Res.*, **47**, 59 ~ 70. [20, 21]

243 **Mackay, T.F.C.** (1985*b*) Transposable element – induced response to artificial selection in *Drosophila melanogaster*. *Genetics*, **111**, 351 ~ 374. [21]

244 **Mackay, T.F.C.** (1989) Mutation and the origin of quantitative variation. pp. 113 ~ 119 in Hill, W.G. and Mackay, T.F.C. (eds), *Evolution and Animal Breeding*. C.A.B. International, Wallingford, UK. [21]

245 **Mackay, T.F.C.** and **Langley, C.H.** (1990) Molecular and phenotypic variance in the *achaete – scute* region of *Drosophila melanogaster*. *Nature*, **348**, 64 ~ 66. [21]

246 **Mackay, T.F.C.**, **Lyman, R.F.** and **Jackson, M.S.** (1992) Effects of *P* element insertions on quantitative traits in *Drosophila melanogaster*. *Genetics*, **130**, 315 ~ 332. [20]

247 **McBride, G.** and **Robertson, A.** (1963) Selection using assortative mating in *Drosophila melanogaster*. *Genet. Res.*, **4**, 356 ~ 369. [10]

248 **McCarthy, J.C.** (1965) Effects of concurrent lactation on litter size and prenatal mortality in an inbred strain of mice. *J. Reprod. Fertil.*, **9**, 29 ~ 39. [14]

249 **McKusick, V.A.** (1978) *Medical Genetic Studies of the Amish: Selected Papers Assembled, with Commentary*. Johns Hopkins Univ. Baltimore, Md., USA. [4]

250 **McKusick, V.A.** (1994) *Mendelian Inheritance in Man*, Johns Hopkins University Press, Baltimore, Md., USA. [21]

251 **McLaren, A.** and **Michie, D.** (1954) Factors affecting vertebral variation in mice. Ⅰ. Variation within an inbred strain. *J. Embryol. Exp. Morph.*, **2**, 149 ~ 160. [15]

252 **McLaren, A.** and **Michie, D.** (1956) Variability of response in experimental animals. *J. Genet.*, **54**, 440 ~ 455. [15]

253 **McMillan, I.** and **Robertson, A.** (1974) The power of methods for the detection of major genes affecting quantitative characters. *Heredity*, **32**, 349 ~ 356. [21]

254 **McNew, R.W.** and **Bell, A.E.** (1976) Comparisons of crossbred and purebred selection for a heterotic trait in highly selected populations of *Tribolium*. *J. Hered.*, **67**, 275 ~ 283. [16]

255 **Mather, K.** (1941) Variation and selection of polygenic characters. *J. Genet.*, **41**, 159 ~ 193. [20]

256 **Mather, K.** (1953) Genetical control of stability in development. *Heredity*, **7**, 297 ~ 336. [15]

257 **Mather, K.** and **Jinks, J.L.** (1977) *Introduction to Biometrical Genetics*. Chapman and Hall, London. [绪论, 8, 16]

258 **Mather, K.** and **Jinks, J.L.** (1982) *Biometrical Genetics*, 3rd edn. Chapman and Hall, London. [绪论, 8, 16, 17]

259 **Mayo, O.** (1987) *The Theory of Plant Breeding*, 2nd edn. Clarendon Press, Oxford, UK. [绪论, 14]

260 **Mayo, O.** (1989) Identification of genes which influence quantitative traits. pp. 141 ~ 146 in Hill, W.G. and Mackay, T.F.C. (eds), *Evolution and Animal Breeding*, C.A.B. International, Wallingford, UK. [21]

261 **Mayo, O.**, **Eckert, S.R.** and **Nugroho, W.H.** (1983) Properties of the major gene index

and related functions. *Human Heredity*, **33**, 205 ~ 212. [21]

262 **Merat, P.** and **Ricard, F.H.** (1974) Etude d'un gène de nanisme lié au sexe chez la poule: Importance de l'état d'engraissement et gain de poids chez l'adulte. *Ann. Génétique Sélection Anim.*, **6**, 211 ~ 217. [21]

263 **Mercer, J.T.** and **Hill, W.G.** (1984) Estimation of genetic parameters for skeletal defects in broiler chickens. *Heredity*, **53**, 193 ~ 203. [18]

264 **Metcalfe, J.A.** and **Turner, J.R.G.** (1971) Gene frequencies in the domestic cats of York: evidence of selection. *Heredity*, **26**, 259 ~ 268. [1]

265 **Meyer, K.** (1989*a*) Restricted Maximum Likelihood to estimate variance components for animal models with several random effects using a derivative – free algorithm. *Génét. Sél. Evol.*, **21**, 317 ~ 340. [10]

266 **Meyer, K.** (1989*b*) Estimation of genetic parameters. pp. 161 ~ 167 in Hill, W.C. and Mackay, T.F.C. (eds), *Evolution and Animal Breeding*, C.A.B. International, Wallingford, UK. [10]

267 **Meyer, K.** and **Hill, W.G.** (1991) Mixed model analysis of a selection experiment for food intake in mice. *Genet. Res.*, **57**, 71 ~ 81. [11]

268 **Mi, M.P.** and **Rashad, M.N.** (1975) Genetic parameters of dermal patterns and ridge counts. *Hum. Hered.*, **25**, 249 ~ 257. [10]

269 **Mikami, H.** and **Fredeen, H.T.** (1979) A genetic study of cryptorchidism and scrotal hernia in pigs. *Can. J. Genet. Cytol.*, **21**, 9 ~ 19. [18]

270 **Minvielle, F.** (1979) Comparing the means of inbred lines with the base population: a model with overdominant loci. *Genet. Res.*, **33**, 89 ~ 92. [14]

271 **Moll, R.H.**, **Lonnquist, J.H.**, **Fortuno, J.V.** and **Johnson, E.C.** (1965) The relationship of heterosis and genetic divergence in maize. *Genetics*, **52**, 139 ~ 144. [14]

272 **Morley, F.H.W.** (1951) Selection for economic characters in Australian Merino sheep. *N.S.W. Dept. Agric. Bull.*, No. 73. [8]

273 **Morley, F.H.W.** (1954) Selection for economic characters in Australian Merino sheep. Ⅳ. The effect of inbreeding. *Austr. J. Agric. Res.*, **5**, 305 ~ 316. [14]

274 **Morton, N.E.** (1955) The inheritance of human birth weight. *Ann. Hum. Genet.*, **20**, 125 ~ 134. [8]

275 **Morton, N.E.** (1974) Analysis of family resemblance. Ⅰ. Introduction. *Am. J. Hum. Genet.*, **26**, 318 ~ 330. [10]

276 **Morton, N.E.** (1978) Effect of inbreeding on IQ and mental retardation. *Proc. Natnl. Acad. Sci., Wash.*, **75**, 3906 ~ 3908. [14]

277 **Morton, N.E.** and **MacLean, C.J.** (1974) Analysis of family resemblance. Ⅲ. Complex segregation of quantitative traits. *Am. J. Hum. Genet.*, **26**, 489 ~ 503. [18, 21]

278 **Mourant, A.E.** (1954) *The Distribution of the Human Blood Groups*. Blackwell, Oxford. [1]

279 **Mousseau, T.A.** and **Roff, D.A.** (1987) Natural selection and the heritability of fitness

components. *Heredity*, **59**, 181 ~ 197. [10, 20]

280 **Muir, W.M.** (1986) Estimation of response to selection and utilization of control populations for additional information and accuracy. *Biometrics*, **42**, 381 ~ 391. [11]

281 **Mukai, T.** (1979) Polygenic mutation. pp. 177 ~ 196 in Thompson, Jr, J.N. and Thoday, J.M. (eds), *Quantitative Genetic Variation*. Academic Press, New York, NY, USA. [20]

282 **Mukai, T.**, **Chigusa, S.I.**, **Mettler, L.E.** and **Crow, J.F.** (1972) Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. *Genetics*, **72**, 335 ~ 355. [20]

283 **Muller, H.J.** and **Oster, I.I.** (1957) Principles of back mutation as observed in *Drosophila* and other organisms. pp. 407 ~ 415 in *Advances in Radiobiology*. *Proc*. *5th Int*. *Conf*. *Radiobiol*. (*Stockh*.), 1956. Oliver and Boyd, Edinburgh. [2]

284 **Nagai, J.**, **Eisen, E.J.**, **Emsley, J.A.B.** and **McAllister, A.J.** (1978) Selection for nursing ability and adult weight in mice. *Genetics*, **88**, 761 ~ 780. [19]

285 **Nagylaki, T.** (1978) The correlation between relatives with assortative mating. *Ann*. *Hum*. *Genet*., **42**, 131 ~ 137. [10]

286 **Nance, W.E.** (1976) Note on the analysis of twin data. *Am*. *J*. *Hum*. *Genet*., **28**, 297 ~ 298. [10]

287 **Nevo, E.** (1978) Genetic variation in natural populations: patterns and theory. *Theor*. *Pop*. *Biol*., **13**, 121 ~ 177. [2]

288 **Nicholas, F.W.** (1980) Size of population required for artificial selection. *Genet*. *Res*., **35**, 85 ~ 105. [12]

289 **Nicholas, F.W.** (1987) *Veterinary Genetics*. Clarendon Press, Oxford, UK. [绪论, 13]

290 **Nordskog, A.W.** (1978) Some statistical properties of an index of multiple traits. *Theor*. *Appl*. *Genet*., **52**, 91 ~ 94. [13]

291 **Oliverio, A.** (1979) Uses of recombinant inbred lines. pp. 197 ~ 218 in Thompson, Jr, J.N. and Thoday, J.M. (eds), *Quantitative Genetic Variation*. Academic Press, New York, NY, USA. [21]

292 **Ollivier, L.** (1974) Optimum replacement rates in animal breeding. *Anim*. *Prod*., **19**, 257 ~ 271. [11]

293 **Orozco, F.** (1976) A dynamic study of genotype – environment interaction with egg laying of *Tribolium castaneum*. *Heredity*, **37**, 157 ~ 171. [19]

294 **Osborne, R.** (1957) The use of sire and dam family averages in increasing the efficiency of selective breeding under a hierarchical mating system. *Heredity*, **11**, 93 ~ 116. [13]

295 **Otsuka, Y.**, **Eberhart, S.A.** and **Russell, W.A.** (1972) Comparisons of prediction formulas for maize hybrids. *Crop Sci*., **12**, 325 ~ 331. [16]

296 **Paterson, A.H.**, **Damon, S.**, **Hewitt, J.D.**, **Zamir, D.**, **Rabinowitch, H.D.**, **Lincoln, S.E.**, **Lander, E.S.** and **Tanksley, S.D.** (1991) Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. *Genetics*, **127**, 181 ~ 197. [21]

297 **Paterson, A.H.**, **DeVerna, J.W.**, **Lanini, B.** and **Tanksley, S.D.** (1990) Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes in an interspecies cross of tomato. *Genetics*, **124**, 735 ~ 742. [21]

298 **Paterson, A.H.**, **Lander, E.S.**, **Hewitt, J.D.**, **Peterson, S.**, **Lincoln, S.E.** and **Tanksley, S.D.** (1988) Resolution of quantitative traits in Mendelian factors using a complete linkage map of restriction fragment length polymorphisms. *Nature*, **335**, 721 ~ 726. [6, 21]

299 **Penrose, L.S.** (1954) Some recent trends in human genetics. *Proc. 9th Internat. Congr. Genet.*, *Caryologia*, **6** (Suppl.), 521 ~ 530. [8]

300 **Perkins, J.M.** and **Jinks, J.L.** (1968) Environmental and genotype – environmental components of variability. Ⅲ. Multiple lines and crosses. *Heredity*, **23**, 339 ~ 356. [8]

301 **Piper, L.R.** and **Bindon, B.M.** (1982) Genetic segregation for fecundity in Booroola Merino sheep. pp. 315 ~ 331 in Barton, R.A. and Smith, W.C. (eds), *Proceedings of the World Congress on Sheep and Beef Cattle Breeding*. Dunmore Press, Palmerston North, Australia. [21]

302 **Piper, L.R.** and **Bindon, B.M.** (1988) The genetics and endocrinology of the Booroola sheep *F* gene. pp. 270 ~ 280 in Weir, B.S., Eisen, E.J., Goodman, M.M. and Namkoong, G. (eds), *Proceedings of the 2nd International Conference on Quantitative Genetics*. Sinauer, Sunderland, Mass., USA. [21]

303 **Piper, L.R.**, **Bindon, B.M.** and **Davis, G.H.** (1985) The single gene inheritance of the high litter size of the Booroola Merino. pp. 115 ~ 125 in Land, R.B. and Robinson, D.W. (eds), *Genetics of Reproduction in Sheep*. Butterworth, London, UK. [21]

304 **Pirchner, F.** (1988) Finding genes affecting quantitative traits in domestic animals. pp. 243 ~ 249 in Weir, B.S., Eisen, E.J., Goodman, M.M. and Namkoong, G. (eds), *Proceedings of the 2nd International Conference on Quantitative Genetics*. Sinauer, Sunderland, Mass., USA. [21]

305 **Pium, M.** (1954) Computation of inbreeding and relationship coefficients. *J. Hered.*, **45**, 92 ~ 94. [5]

306 **Ponzoni, R.W.** and **James, J.W.** (1978) Possible biases in heritability estimates from intra – class correlation. *Theor. Appl. Genet.*, **53**, 25 ~ 27. [10]

307 **Pooni, H.S.** and **Jinks, J.L.** (1979) Sources and biases of the predictors of the properties of recombinant inbreds produced by single seed descent. *Heredity*, **42**, 41 ~ 48. [16]

308 **Pooni, H.S.**, **Jinks, J.L.** and **Jayasekara, N.E.M.** (1978) An investigation of gene action and genotype × environment interaction in two crosses of *Nicotiana rustica* by triple test cross and inbred line analysis. *Heredity*, **43**, 83 ~ 91. [9]

309 **Pooni, H.S.**, **Jinks, J.L.** and **Singh, R.K.** (1984) Methods of analysis and the estimation of the genetic parameters from a diallel set of crosses. *Heredity*, **52**, 243 ~ 253. [15]

310 **Powers, L.** (1950) Determining scales and the use of transformations in studies on weight per locule of tomato fruit. *Biometrics*, **6**, 145 ~ 163. [17]

311 **Powers, L.** (1952) Gene recombination and heterosis. pp.298 ~ 319 in Gowen, J.W. (ed.), *Heterosis*. Iowa State Coll., Ames, Iowa, USA. [14]

312 **Preston, T.R.** and **Willis, M.B.** (1970) *Intensive Beef Production*. Pergamon, Oxford. [10]

313 **Price, G.R.** (1972) Fisher's 'fundamental theorem' made clear. *Ann. Hum. Genet.*, **36**, 129 ~ 140. [20]

314 **Prout, T.** (1962) The effects of stabilizing selection on the time of development in *Drosophila melanogaster*. *Genet. Res.*, **3**, 364 ~ 382. [20]

315 **Prout, T.** (1965) The estimation of fitnesses from genotypic frequencies. *Evolution*, **19**, 546 ~ 551. [2]

316 **Raine, D.N.**, **Cooke, J.R.**, **Andrews, W.A.** and **Mahon, D. F.** (1972) Screening for inherited metabolic disease by plasma chromatography (Scriver) in a large city. *Brit. Med. J.*, **3**, 7 ~ 13. [1]

317 **Rao, D.C.**, **Morton, N.E.** and **Yee, S.** (1974) Analysis of family resemblance. Ⅱ. A linear model for familial correlation. *Am. J. Hum. Genet.*, **26**, 331 ~ 359. [10]

318 **Rao, D.C.**, **Morton, N.E.** and **Yee, S.** (1976) Resolution of cultural and biological inheritance by path analysis. *Am. J. Hum. Genet.*, **28**, 228 ~ 242. [10]

319 **Rasmuson, M.** (1952) Variation in bristle number of *Drosophila melanogaster*. *Acta Zoolog., Stockh.*, **33**, 277 ~ 307. [15]

320 **Reeve, E.C.R.** (1955*a*) Inbreeding with homozygotes at a disadvantage. *Ann. Hum. Genet.*, **19**, 332 ~ 346. [5]

321 **Reeve, E.C.R.** (1955*b*) The variance of the genetic correlation coefficient. *Biometrics*, **11**, 357 ~ 374. [19]

322 **Reeve, E.C.R.** (1961) A note on non - random mating in progeny tests. *Genet. Res.*, **2**, 195 ~ 203. [10]

323 **Reeve, E.C.R.** and **Robertson, F.W.** (1953) Studies in quantitative inheritance. Ⅱ. Analysis of a strain of *Drosophila melanogaster* selected for long wings. *J. Genet.*, **51**, 276 ~ 316. [10, 19]

324 **Reeve, E.C.R.** and **Robertson, F.W.** (1954) Studies in quantitative inheritance. Ⅵ. Sternite chaeta number in *Drosophila*: a metameric quantitative character. *Z. indukt. Abstam. - u. Vererblehre*, **86**, 269 ~ 288. [8]

325 **Reich, T.**, **James, J.W.** and **Morris, C.A.** (1972) The use of multiple thresholds in determining the mode of transmission of semi - continuous traits. *Ann. Hum. Genet.*, **36**, 163 ~ 184. [18]

326 **Rendel, J.M.** and **Robertson, A.** (1950) Estimation of genetic gain in milk yield by selection in a closed herd of dairy cattle. *J. Genet.*, **50**, 1 ~ 8. [11]

327 **Rendel, J.M.**, **Robertson, A.**, **Asker, A.A.**, **Khishin, S.S.** and **Ragab, M.T.** (1957) The inheritance of milk production characteristics. *J. Agric. Sci.*, **48**, 426 ~ 432. [8]

328 **Rich, S.S.**, **Bell, A.E.** and **Wilson, S.P.** (1979) Genetic drift in small populations of *Tribolium*. *Evolution*, **33**, 579 ~ 584. [3]

329 **Roberts, D.F.**, **Billewicz, W.Z.** and **McGregor, I.A.** (1978) Heritability of stature in a West African population. *Ann. Hum. Genet.*, **42**, 15 ~ 24. [10]

330 **Roberts, R.C.** (1960) The effects on litter size of crossing lines of mice inbred without selection. *Genet. Res.*, **1**, 239 ~ 253. [14]

331 **Roberts, R.C.** (1966*a*) The limits to artificial selection for body weight in the mouse. Ⅰ. The limits attained in earlier experiments. *Genet. Res.*, **8**, 347 ~ 360. [12]

332 **Roberts, R.C.** (1966*b*) The limits to artificial selection for body weight in the mouse. Ⅱ. The genetic nature of the limits. *Genet. Res.*, **8**, 361 ~ 375. [12]

333 **Roberts, R.C.** (1967*a*) The limits to artificial selection for body weight in the mouse. Ⅲ. Selection from crosses between previously selected lines. *Genet. Res.*, **9**, 73 ~ 85. [12]

334 **Roberts, R.C.** (1967*b*) The limits to artificial selection for body weight in the mouse. Ⅳ. Sources of new genetic variance -irradiation and outcrossing. *Genet. Res.*, **9**, 87 ~ 98. [12]

335 **Roberts, R.C.** and **Smith, C.** (1982) Genes with large effects. Theoretical aspects in livestock breeding. *Proceedings of the 2nd World Congress on Genetics Applied to Livestock Production*, Vol.6, 420 ~ 438. [21]

336 **Robertson, A.** (1952) The effect of inbreeding on the variation due to recessive genes. *Genetics*, **37**, 189 ~ 207. [15]

337 **Robertson, A.** (1954) Inbreeding and performance in British Friesian cattle. *Proc. Brit. Soc. Anim. Prod.*, 1954, 87 ~ 92. [14]

338 **Robertson, A.** (1956) The effect of selection against extreme deviants based on deviation or on homozygosis. *J. Genet.*, **54**, 236 ~ 248. [20]

339 **Robertson, A.** (1959*a*) Experimental design in the evaluation of genetic parameters. *Biometrics*, **15**, 219 ~ 226. [10]

340 **Robertson, A.** (1959*b*) The sampling variance of the genetic correlation coefficient. *Biometrics*, **15**, 469 ~ 485. [19]

341 **Robertson, A.** (1960) A theory of limits in artificial selection. *Proc. Roy. Soc. London, B.*, **153**, 234 ~ 249. [12]

342 **Robertson, A.** (1962) Selection for heterozygotes in small populations. *Genetics*, **47**, 1291 ~ 1300. [4]

343 **Robertson, A.** (1965) The interpretation of genotypic ratios in domestic animal populations. *Anim. Prod.*, **7**, 319 ~ 324. [3]

344 **Robertson, A.** (1966) A mathematical model of the culling process in dairy cattle. *Anim. Prod.*, **8**, 95 ~ 108. [20]

345 **Robertson, A.** (1967) The nature of quantitative genetic variation. pp.265 ~ 280 in A. Brink (ed.), *Heritage From Mendel*. The University of Wisconsin Press, Madison, Wis., USA. [21]

346 **Robertson, A.** (1970) A theory of limits in artificial selection with many linked loci. pp. 246 ~ 288 in Kojima, K. (ed.), *Mathematical Topics in Population Genetics*, Vol. 1, Springer – Verlag, Berlin. [12]

347 **Robertson, A.** (1977*a*) The effect of selection on the estimation of genetic parameters. *Z. Tierzüchtg. Züchtgsbiol.*, **94**, 131 ~ 135. [10, 11]

348 **Robertson, A.** (1977*b*) The non – linearity of offspring – parent regression. pp. 297 ~ 304 in Pollak, E., Kempthorne, O. and Bailey, T.B. (eds), *Proc. Int. Conf. Quantitative Genetics*, Iowa State Univ., Ames, Iowa, USA. [11, 12]

349 **Robertson, D.S.** (1985) A possible technique for isolating genic DNA for quantitative traits in plants. *J. Theor. Biol.*, **117**, 1 ~ 10. [21]

350 **Robertson, F.W.** (1955) Selection response and the properties of genetic variation. *Cold Spring Harbor Symp. Quant. Biol.*, **20**, 166 ~ 177. [12]

351 **Robertson, F.W.** (1957*a*) Studies in quantitative inheritance. Ⅹ. Genetic variation of ovary size in *Drosophila*. *J. Genet.*, **55**, 410 ~ 427. [8, 10]

352 **Robertson, F.W.** (1957*b*) Studies in quantitative inheritance. Ⅺ. Genetic and environmental correlation between body size and egg production in *Drosophila melanogaster*. *J. Genet.*, **55**, 428 ~ 443. [8, 10]

353 **Robertson, F.W.** and **Reeve, E.C.R.** (1952) Heterozygosity, environmental variation and heterosis. *Nature*, **170**, 296. [15]

354 **Robinson, P.** and **Bray, D.F.** (1965) Expected effects on the inbreeding coefficient and rate of gene loss of four methods of reproducing finite diploid populations. *Biometrics*, **21**, 447 ~ 458. [4]

355 **Robson, E.B.** (1955) Birth weight in cousins. *Ann. Hum. Genet.*, **19**, 262 ~ 268. [8]

356 **Roff, D.** (1981) On being the right size. *Amer. Nat.*, **118**, 405 ~ 422. [20]

357 **Roff, D.A.** and **Mousseau, T.A.** (1987) Quantitative genetics and fitness: lessons from *Drosophila*. *Heredity*, **58**, 103 ~ 118. [10, 20]

358 **Rollins, W.C.**, **Tanaka, M.**, **Nott, C.F.G.** and **Theissen, R.B.** (1972) On the mode of inheritance of double – muscled conformation in bovines. *Hilgardia*, **41**, 433 ~ 456. [21]

359 **Rose, M.R.** and **Charlesworth, B.** (1981) Genetics of life history in *Drosophila melanogaster*. Ⅰ. Sib analysis of adult females. *Genetics*, **97**, 173 ~ 186. [20]

360 **Rumball, W.**, **Franklin, I.R.**, **Frankham, R.** and **Sheldon, B.L.** (1994) Decline in heterozygosity under full – sib and double first – cousin iubreeding in *Drosophila melanogaster*. *Genetics*, **136**, 1039 ~ 1049. [5]

361 **Russell, E.S.** (1949) A quantitative histological study of the pigment found in the coat – color mutants of the house mouse. Ⅳ. The nature of the effects of genic substitution in five major allelic series. *Genetics*, **34**, 146 ~ 166. [7]

362 **Rutledge, J.J.**, **Eisen, E.J.** and **Legates, J.E.** (1973) An experimental evaluation of genetic correlation. *Genetics*, **75**, 709 ~ 726. [10, 19]

363 **Sales, J.** and **Hill, W.G.** (1976) Effect of sampling errors on efficiency of selection in-

dices. 2. Use of information on associated traits for improvement of a single important trait. *Anim. Prod.*, **23**, 1 ~ 14. [19]

364 **Sax, K.** (1923) The association of size differences with seed – coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics*, **8**, 522 ~ 560. [21]

365 **Schäfer, W.** (1937) Über die Zunahme der Isozygotie (Gleicherbarkeit) bei fortgesetzter Bruder – Schwester – Inzucht. *Z. indukt. Abstamm. – u. Vererblehre*, **72**, 50 ~ 79. [5]

366 **Schlager, G.** and **Dickie, M.M.** (1971) Natural mutation rates in the house mouse. Estimates for five specific loci and dominant mutations. *Mutation Res.*, **11**, 89 ~ 96. [2]

367 **Schluter, D.**, **Price, T.D.** and **Grant, P.R.** (1985) Ecological character displacement in Darwin's finches. *Science*, **227**, 1056 ~ 1058. [20]

368 **Schull, W.J.** (1962) Inbreeding and maternal effects in the Japanese. *Eugen. Quart.*, **9**, 14 ~ 22. [14]

369 **Searle, A.G.** (1949) Gene frequencies in London's cats. *J. Genet.*, **49**, 214 ~ 220. [1]

370 **Searle, S.R.** (1965) The value of indirect selection: Ⅰ. Mass selection. *Biometrics*, **21**, 682 ~ 707. [19]

371 **Searle, S.R.** (1971) *Linear Models*. Wiley, New York. [10]

372 **Shank, D.B.** and **Adams, M.W.** (1960) Environmental variability within inbred lines and single crosses of maize. *J. Genet.*, **57**, 119 ~ 126. [15]

373 **Shaw, R.G.** (1987) Maximum – likelihood approaches applied to quantitative genetics of natural populations. *Evolution*, **41**, 812 ~ 826. [10, 11, 12]

374 **Sheridan, A.K.** (1981) Crossbreeding and heterosis. *Anim. Breed. Abstr.*, **49**, 131 ~ 144. [16]

375 **Sheridan, A.K.** and **Barker, J.S.F.** (1974) Two – trait selection and the genetic correlation. Ⅱ. Changes in the genetic correlation during two – trait selection. *Austr. J. Bio. Sci.*, **27**, 89 ~ 101. [19]

376 **Sheridan, A.K.**, **Frankham, R.**, **Jones, L.P.**, **Rathie, K.A.** and **Barker, J.S.F.** (1968) Partitioning of variance and estimation of genetic parameters for various bristle number characters of *Drosophila melanogaster*. *Theor. Appl. Genet.*, **38**, 179 ~ 187. [19]

377 **Shrimpton, A.E.** and **Robertson, A.** (1988*a*) The isolation of polygenic factors controlling bristle score in *Drosophila melanogaster*. I. Allocation of third chromosome bristle effects to chromosome sections. *Genetics*, **118**, 437 ~ 443. [21]

378 **Shrimpton, A.E.** and **Robertson, A.** (1988*b*) The isolation of polygenic factors controlling bristle score in *Drosophila melanogaster*. 2. Distribution of third chromosome bristle effects within chromosome sections. *Genetics*, **118**, 445 ~ 459. [6, 21]

379 **Sidwell, G.M.**, **Everson, D.O.** and **Terrill, C.E.** (1962) Fertility, prolificacy and lamb livability of some pure breeds and their crosses. *J. Anim. Sci.*, **21**, 875 ~ 879. [16]

380 **Sidwell, G.M.**, **Everson, D.O.** and **Terrill, C.E.** (1964) Lamb weights in some pure breeds and crosses. *J. Anim. Sci.*, **23**, 105 ~ 110. [16]

381 **Simmonds, N.W.** (1979) *Principles of Crop Improvement*. Longman, London. [6, 15, 16]

382 **Sing, C.F., Boerwinkle, E., Moll, P.P.** and **Templeton, A.R.** (1988) Characterization of genes affecting quantitative traits in humans. pp.250 ~ 269 in Weir, B.S., Eisen, E.J., Goodman, M.M. and Namkoong, G. (eds), *Proceedings of the 2nd International Conference on Quantitative Genetics*. Sinauer, Sunderland, Mass., USA. [21]

383 **Skibinski, D.O.F., Woodwark, M.** and **Ward, R.D.** (1993) A quantitative test of the neutral theory using pooled allozyme data. *Genetics*, **135**, 233 ~ 248. [4]

384 **Smith, C.** (1962) Estimation of genetic change in farm livestock using field records. *Anim. Prod.*, **4**, 239 ~ 251. [11]

385 **Smith, C.** (1984) Rates of genetic change in farm livestock. *Res. and Devel. in Agric.*, **1**, 79 ~ 85. [6, 12, 13]

386 **Smith, C.** (1988) Potential for animal breeding, current and future. Ch.13 in Weir, B.S., Eisen, E.J., Goodman, M.M. and Namkoong, G. (eds), *Proc. 2nd Internat. Conf. on Quantitative Genetics*. Sinauer, Sunderland, Mass., USA. [6, 12, 13]

387 **Smith, C.** and **Bampton, P.R.** (1977) Inheritance of reaction to halothane anaesthesia in pigs. *Genet. Res.*, **29**, 287 ~ 292. [21]

388 **Smith, C., King, J.W.B.** and **Gilbert, N.** (1962) Genetic parameters of British Large White pigs. *Anim. Prod.*, **4**, 128 ~ 143. [10, 19]

389 **Smith, C.A.B.** (1970) A note on testing the Hardy – Weinberg Law. *Ann. Hum. Genet.*, **33**, 377 ~ 383. [1]

390 **Smith, H.H.** (1952) Fixing transgressive vigor in *Nicotiana rustica*. pp.161 ~ 174 in Gowen, J.W. (ed.), *Heterosis*, Iowa State Coll., Ames, Iowa, USA. [14]

391 **Snedecor, G.W.** and **Cochran, W.G.** (1967) *Statistical Methods*. (6th edn). Iowa State Univ., Ames, Iowa, USA. [10]

392 **Sokal, R.R.** and **Rohlf, F.J.** (1981) *Biometry*. Freeman, New York, NY, USA. [21]

393 **Soller, M.** and **Beckmann, J.S.** (1988) Genomic genetics and the utilization for breeding purposes of genetic variation between populations. pp.161 ~ 188 in Weir, B.S., Eisen, E.J., Goodman, M.M. and Namkoong, G. (eds), *Proceedings of the 2nd International Conference on Quantitative Genetics*. Sinauer, Sunderland, Mass., USA. [21]

394 **Soller, M.** and **Beckmann, J.S.** (1990) Marker – based mapping of quantitative trait loci using replicated progenies. *Theor. Appl. Genet.*, **80**, 205 ~ 208. [21]

395 **Soller, M., Brody, T.** and **Genizi, A.** (1976) On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theor. Appl. Genet.*, **47**, 35 ~ 39. [21]

396 **Spickett, S.G.** and **Thoday, J.M.** (1966) Regular responses to selection. Ⅲ. Interaction between located polygenes. *Genet. Res.*, **7**, 96 ~ 121. [21]

397 **Spiers, J.G.C.** (1974) The effects of larval competition on a quantitative character in *Drosophila melanogaster*. PhD Thesis, University of Edinhurgh. [20]

398 **Sprague, G.F.** and **Tatum, L.A.** (1942) General *vs*. specific combining ability in single crosses of corn. *J. Amer. Soc. Agron.*, **34**, 923 ~ 932. [15]

399 **Stam, P.** (1980) The distribution of the fraction of the genome identical by descent in finite random mating populations. *Genet. Res.*, **35**, 131 ~ 155. [5]

400 **Stam, P.** and **Zeven, A.C.** (1981) The theoretical proportion of the donor genome in nearisogenic lines of self – fertilizers bred by backcrossing. *Euphytica*, **30**, 227 ~ 238. [5]

401 **Stern, C.** (1973) *Principles of Human Genetics* (3rd edn). Freeman, San Francisco. [2,10]

402 **Stockard, C.R.** (1941) *The Genetic and Endocrine Basis for Differences in Form and Behaviour*. Wistar Institute, Philadelphia. [12]

403 **Strang, G.S.** and **Smith, C.** (1979) A note on the heritability of litter traits in pigs. *Anim. Prod.*, **28**, 403 ~ 406. [10]

404 **Stuber, C.W., Lincoln, S.E., Wolff, D.W., Helentjaris, T.** and **Lander, E.S.** (1992) Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics*, **132**, 823 ~ 839. [21]

405 **Sved, J.A.** (1971) *An estimate of heterosis in Drosophila melanogaster*. *Genet. Res.*, **18**, 97 ~ 105. [20]

406 **Sved, J.A.** (1975) Fitness of third chromosome homozygotes in *Drosophila melanogaster*. *Genet. Res.*, **25**, 197 ~ 200. [20]

407 **Sved, J.A.** (1989) The measurement of fitness in *Drosophila*. pp.99 ~ 104 in Hill, W.G. and Mackay, T.F.C. (eds), *Evolution and Animal Breeding*. C.A.B. International, Wallingford, UK. [20]

408 **Takahashi, J.S., Pinto, L.H.** and **Vitaterna, M.H.** (1994) Forward and reverse genetic approaches to behavior in the mouse. *Science*, **264**, 1724 ~ 1733. [21]

409 **Tanksley, S.D.** (1993) Mapping polygenes. *Annu. Rev. Genet.*, **27**, 205 ~ 233. [21]

410 **Tantawy, A.O.** and **Reeve, E.C.R.** (1956) Studies in quantitative inheritance. IX. The effects of inhreeding at different rates in *Drosophila melanogaster*. *Z. indukt. Abstamm. – u. Vererblehre*, **87**, 648 ~ 667. [14, 15]

411 **Thoday, J.M.** (1961) Location of polygenes. *Nature*, **191**, 368 ~ 370. [21]

412 **Thoday, J.M.** (1979) Polygene mapping: Uses and limitations. pp.220 ~ 233 in Thompson, Jr, J.N. and Thoday, J.M. (eds), *Quantitative Genetic Variation*. Academic Press, New York, NY, USA. [21]

413 **Thompson, J.N.Jr** (1975) A test of the influence of isoallelic variation upon a quantitative character. *Heredity*, **35**, 401 ~ 406. [21]

414 **Thompson, J.N.Jr** and **Thoday, J.M.** (1974) A definition and standard nomenclature for 'polygenic loci'. *Heredity*, **33**, 430 ~ 437. [21]

415 **Thompson, R.** (1976) The estimation of maternal genetic variances. *Biometrics*, **32**, 903 ~ 917. [9]

416 **Thompson, R.** (1979) Sire evaluation. *Biometrics*, **35**, 339 ~ 353. [13]

417 **Thompson, R.** (1982) Methods of estimation of genetic parameters. pp. 95 ~ 103 in *2nd World Congress on Genetics Applied to Livestock Production*, Vol. 5. Editorial Garsi, Madrid, Spain. [10]

418 **Thompson, R.** (1989) Design of experiments to estimate genetic parameters within populations. pp. 169 ~ 174 in Hill, W. G. and Mackay, T. F. C. (eds), *Evolution and Animal Breeding*. C. A. B. International, Wallingford, UK. [13]

419 **Turelli, M.** (1984) Heritable genetic variation via mutation – selection balance: Lerch's zeta meets the abdominal bristle. *Theor. Pop. Biol.*, **25**, 138 ~ 193. [20]

420 **Turelli, M.** (1985) Effects of pleiotropy on predictions concerning mutation – selection balance for polygenic traits. *Genetics*, **111**, 165 ~ 195. [20]

421 **Turner, H. N.** and **Young, S. Y.** (1969) *Quantitative Genetics in Sheep Breeding*. Macmillan, Melbourne, Australia. [10]

422 **Van Vleck, L. D.** (1993) *Selection Index and Introduction to Mixed Model Methods*. C. R. C. Press, Boca Raton, Fla., USA. [13]

423 **Van Vleck, L. D.** and **Henderson, C. R.** (1961) Empirical sampling estimates of genetic correlations. *Biometrics*, **17**, 359 ~ 371. [19]

424 **Waddington, C. H.** (1942) Canalisation of development and the inheritance of acquired characters. *Nature*, **150**, 563. [18]

425 **Waddington, C. H.** (1953) Genetic assimilation of an acquired character. *Evolution*, **7**, 118 ~ 126. [18]

426 **Waddington, C. H.** (1957) *The Strategy of the Genes*. Allen and Unwin, London. [20]

427 **Wallace, B.** (1958) The comparison of observed and calculated zygotic distributions. *Evolution*, **12**, 113 ~ 115. [1]

428 **Wallace, B.** (1963) The elimination of an autosomal lethal from an experimental population of *Drosophila melanogaster*. *Am. Nat.*, **97**, 65 ~ 66. [2]

429 **Wallace, B.** (1968) *Topics in Population Genetics*. Norton, New York. [1, 2]

430 **Wallace, B.** and **Vetukhiv, M.** (1955) Adaptive organization of the gene pools of *Drosophila* populations. *Cold Spring Harbor Symp. Quant. Biol.*, **20**, 303 ~ 309. [14]

431 **Waller, J. H.** (1971) Differential reproduction: its relation to IQ test score, education, and occupation. *Social Biol.*, **18**, 122 ~ 136. [20]

432 **Warwick, E. J.** and **Lewis, W. L.** (1954) Increase in frequency of a deleterious recessive gene in mice. *J. Hered.*, **45**, 143 ~ 145. [7]

433 **Weatherspoon, J. H.** (1970) Comparative yields of single, three – way, and double crosses of maize. *Crop Sci.*, **10**, 157 ~ 159. [16]

434 **Webb, A. J.**, **Carden, A. E.**, **Smith, C.** and **Imlah, P.** (1982) Porcine stress syndrome in pig breeding. *Proceedings of the 2nd World Congress on Genetics Applied to Livestock Production*, Vol. 5, 588 ~ 608. [21]

435 **Weinberg, W.** (1908) Über den Nachweis der Vererbung beim Menschen. *Jh. Ver. vaterl Naturk. Württemb.*, **64**, 369 ~ 382. [1]

436 **Weir, B.S.** (1990) *Genetic Data Analysis*. Sinauer, Sunderland, Mass., USA. [绪论, 1]

437 **Weir, B.S.** and **Cockerham, C.C.** (1977) Two - locus theory in quantitative genetics. pp.247 ~ 269 in Pollak, E., Kempthorne, O. and Bailey, T.B. (eds), *Proc. Int. Conf. Quantitative Genetics*, Iowa State Univ., Ames, Iowa, USA. [15]

438 **Weir, B.S.** and **Cockerham, C.C.** (1979) Estimation of linkage disequilibrium in randomly mating populations. *Heredity*, **42**, 105 ~ 111. [1]

439 **Weir, B.S.**, **Cockerham, C.C.** and **Reynolds, J.** (1980) The effects of linkage and linkage disequilibrium on the covariances of non - inbred relatives. *Heredity*, **45**, 351 ~ 359. [9]

440 **Weller, J.I.**, **Soller, M.** and **Brody, T.** (1988) Linkage analysis of quantitative traits in an interspecific cross of tomato (*Lycopersicon esculentum* × *Lycopersicon pimpinellifolium*) by means of genetic markers. *Genetics*, **118**, 329 ~ 339. [21]

441 **White, J.M.** (1972) Inbreeding effects upon growth and maternal ability in laboratory mice. *Genetics*, **70**, 307 ~ 317. [14]

442 **Willham, R.L.** (1963) The covariance between relatives for characters composed of components contributed by related individuals. *Biometrics*, **19**, 18 ~ 27. [9]

443 **Willham, R.** (1972) The role of maternal effects in animal breeding: Ⅲ. Biometrical aspects of maternal effects in animals. *J. Anim. Sci.*, **35**, 1288 ~ 1293. [9]

444 **Wills, C.** (1978) Rank - order selection is capable of maintaining all genetic polymorphisms. *Genetics*, **89**, 403 ~ 417. [2]

445 **Wolstenholme, D.R.** and **Thoday, J.M.** (1963) Effects of disruptive selection. Ⅶ. A third chromosome polymorphism. *Heredity*, **10**, 413 ~ 431. [21]

446 **Wright, A.J.** (1985) Diallel designs, analyses, and reference populations. *Heredity*, **54**, 307 ~ 311. [15]

447 **Wright, S.** (1921) Systems of mating. *Genetics*, **6**, 111 ~ 178. [绪论, 10]

448 **Wright, S.** (1933) Inbreeding and homozygosis. *Proc. Natnl. Acad. Sci. Wash.*, **19**, 411 ~ 420. [5]

449 **Wright, S.** (1942) Statistical genetics and evolution. *Bull. Amer. Math. Soc.*, **48**, 223 ~ 246. [4]

450 **Wright, S.** (1951) The genetical structure of populations. *Ann. Eugen. Lond.*, **15**, 323 ~ 354. [4]

451 **Wright, S.** (1952) The theoretical variance within and among subdivisions of a population that is in a steady state. *Genetics*, **37**, 312 ~ 321. [3]

452 **Wright, S.** (1952) The genetics of quantitative variability. pp.5 ~ 14 in Reeve, E.C.R. and Waddington, C.H. (eds), *Quantitative Inheritance*. HMSO, London, UK. [21]

453 **Wright, S.** (1968) *Evolution and the Genetics of Populations*. Vol.1, *Genetic and Biometric Foundations*. Univ. of Chicago, Chicago. [绪论, 8, 17]

454 **Wright, S.** (1969) *Evolution and the Genetics of Populations*. Vol.2, *The Theory of Gene Frequencies*. Univ. of Chicago, Chicago. [绪论, 4, 5, 15, 20]

455 **Wright, S.** (1977) *Evolution and the Genetics of Populations*. Vol.3, *Experimental Results and Evolutionary Deductions*. Univ. of Chicago, Chicago. [绪论, 15]

456 **Wright, S.** (1978) *Evolution and the Genetics of Populations*. Vol.4, *Variability within and among Natural Populations*. Univ. of Chicago, Chicago. [绪论]

457 **Wright, S.** and **Kerr, W.E.** (1954) Experimental studies of the distribution of gene frequencies in very small populations of *Drosophila melanogaster*. Ⅱ. Bar. *Evolution*, **8**, 225 ~ 240. [4]

458 **Wright, S.** and **McPhee, H.C.** (1925) An approximate method of calculating coefficients of inbreeding and relationship from livestock pedigrees. *J. Agric. Res.*, **31**, 377 ~ 383. [5]

459 **Yamada, Y.** (1962) Genotype by environment interaction and genetic correlation of the same trait under different environments. *Jap. J. Genet.*, **37**, 498 ~ 509. [19]

460 **Yamada, Y.** and **Bell, A.E.** (1969) Selection for larval growth in *Tribolium* under two levels of nutrition. *Genet. Res.*, **13**, 175 ~ 195. [19]

461 **Yamada, Y., Yokouchi, K.** and **Nishida, A.** (1975) Selection index when genetic gains of individual traits are of primary concern. *Jap. J. Genet.*, **50**, 33 ~ 41. [19]

462 **Yoo, B.H.** (1980*a*) Long – term selection for a quantitative character in large replicate populations of *Drosophila melanogaster*. Ⅰ. Responses to selection. *Genet. Res.*, **35**, 1 ~ 17. [12]

463 **Yoo, B.H.** (1980*b*) Long – term selection for a quantitative character in large replicate populations of *Drosophila melanogaster*. Ⅱ. Lethals and visible mutants with large effects. *Genet. Res.*, **35**, 19 ~ 31. [12]

464 **Yoo, B.H.** (1980*c*) Long – term selection for a quantitative character in large replicate populations of *Drosophila melanogaster*. Ⅲ. The nature of residual genetic variability. *Theor. Appl. Genet.*, **57**, 25 ~ 32. [12]

465 **Yoon, C.H.** (1955) Homeostasis associated with heterozygosity in the genetics of time of vaginal opening in the house mouse. *Genetics*, **40**, 297 ~ 309. [15]

466 **Zeleny, C.** (1922) The effect of selection for eye facet number in the white bar – eye race of *Drosophila melanogaster*. *Genetics*, **7**, 1 ~ 115. [6]

467 **Zeng, Z.-B.** (1994) Precision mapping of quantitative trait loci. *Genetics*, **136**, 1457 ~ 1468. [21]

468 **Zuberi, M.I.** and **Gale, J.S.** (1976) Variation in wild populations of *Papaver dubium*. *Heredity*, **36**, 359 ~ 368. [8]

[General Information]
书名=数量遗传学导论 （第四版）
作者=
页数=377
SS号=11108946
出版日期=

封面 -23
书名 -21
版权 -20
前言 -19
目录 -8
目 录中文版序言（一）中文版序言（二）译者的话第四版序言第三版序言鸣谢绪论第1章群体的遗传组成 1
基因和基因型的频率 1
自然群体中的孟德尔变异 3
变化的原因 3
哈迪—温伯格平衡 4
哈迪—温伯格法则 4
哈迪—温伯格法则的应用 6
交配频率和哈迪—温伯格法则的另一种证明 9
复等位基因 10
性连锁基因 11
多于一个位点 13
非随机交配 16
选型交配 16
习题 17
第2章基因频率的变化 19
迁移 19
突变 19
非频发突变 19
频发突变 20
选择 21
选择下基因频率的变化 22
选择的有效性 24
需要的世代数 26
平均适合度和负荷 27
平衡 27
突变和选择之间的平衡 27
平衡的改变 29
有利于杂合体的选择 30
多态性 33
习 题 36
第3章小群体：Ⅰ．简化条件下基因频率的变化 38
理想群体 39
抽样 40
基因频率的方差 40
固 定 43
基因型频率 44
近 交 45
理想群体中的近交 46
基因频率的方差 48
基因型频率 48
习题 50
第4章小群体：Ⅱ．较少简化条件 52
有效群体大小 52

极近亲交配的排除 52
雄性和雌性数目不同 53
连续世代中数目不等 53
家系大小的非随机分布 54
最低程度的近交 55
世代重叠 55
突变、迁移和选择 57
非频发中性突变 57
频发突变和迁移 58
选择 60
自然群体中的随机漂变 60
多态性 61
中性理论 62
习题 63
第5章小群体：Ⅲ．系谱群体和近亲近交 65
系谱群体 65
一个个体的近交系数 65
共祖率或亲缘关系 67
规则的近交系统 69
近亲近交 70
固 定 71
反复回交 72
杂交和后续世代 73
混合的近交和杂交 74
基础的改变：结构化群体 75
突 变 76
有利于杂合体的选择 76
习题 77
第6章连续变异 79
数量性状 80
数量性状的特性 82
习题 84
第7章数值与均数 86
群体平均值 87
平均效应 89
育种值 91
显性离差 92
互作离差 94
习题 95
第8章方差 97
方差组分 97
作为总方差部分的组分 97
遗传决定程度VG/VP的估计 98
方差的遗传组分 99
加性和显性方差 99
总遗传方差 100
互作方差 103
由不平衡引起的方差 103
基因型和环境之间的相关与互作 104

相 关 104
互 作 104
环境方差 106
多次度量：重复力 108
方差分剖的总结 113
习题 113
第9章亲属间的相似性 115
遗传协方差 116
后代与一个亲本 116
后代与中亲值 117
半同胞 118
全同胞 120
孪生子 121
概 要 121
上位互作 122
环境协方差 123
表型相似性 125
习题 125
第10章遗传力 127
遗传力的估计 129
亲子回归 130
同胞分析 131
后代对母畜的公畜内回归 134
联合估计值 134
双胞胎和人类数据 135
选型交配 137
估计值的精确度和实验设计 140
亲子回归 140
同胞分析 142
亲本选择 142
习题 143
第11章选择：Ⅰ．反应及其预测 145
选择反应 146
反应的预测 147
选择差和选择强度 148
反应的改良 150
反应的度量 152
世代平均值的变异性 152
加权选择差 154
现实遗传力 155
母体效应 156
人工选择下基因频率的变化 156
选择对方差的影响 158
习题 161
第12章选择：Ⅱ．实验结果 164
短期结果 164
反应的重复性 164
抽样方差 165
反应的不对称性 166

长期结果 169
选择极限 169
突 变 173
选择极限的原因 175
位点（有效因子）数目和标准化效应 176
习题 177
第13章选择：Ⅲ．来自亲属的信息 179
选择标准 180
简单方法 180
反应的预测 182
合并选择 184
各种方法的相对优点 186
指数选择 188
指数的构建 188
准确度 190
选择反应 191
实际成就 192
习题 192
第14章近交和杂交：Ⅰ．平均值的变化 194
近交衰退 194
选择的效应 197
杂种优势 199
单交 200
习题 205
第15章近交和杂交：Ⅱ．方差的改变 207
近交 207
遗传方差的再分配 207
环境方差 210
近交系的一致性 211
突变 212
亚系歧化 213
杂 交 214
杂种之间的方差 214
配合力 215
习题 219
第16章近交和杂交：Ⅲ．应用 222
配合力的选择 223
三元和四元杂交；回交 224
正反交反复选择（RRS） 226
超显性 226
自然自体受精植物 227
习题 227
第17章尺度 229
分布与方差 229
互作 232
结论 234
习题 234
第18章阈性状 236
趋向和阈值 236

两种类别，一个阈值 236
趋向模型的充分性 239
尺度关系 240
三种类别，两个阈值 240
阈性状的选择 243
习题 245
第19章相关性状 247
遗传和环境相关 247
遗传相关的估计 249
选择的相关反应 250
间接选择 252
基因型—环境互作 254
指数选择 257
指数的构建 257
反 应 259
选择对遗传相关的影响 262
习题 263
第20章 自然选择下的数量性状 265
自然选择 265
适合度及其组分 265
数量性状和适合度之间的关系 267
“适合度轮廓” 267
对自然选择的反应 268
适合度 268
相关反应 268
选择的强度 269
平衡群体 270
适合度 270
主要组分 271
具有中间值最优的性状 272
具有中间值最小适合度的性状 274
中性性状 274
突变引起的变异来源 275
突变方差 275
遗传变异的保持 276
中性突变和随机漂变之间的平衡 277
突变—选择平衡 277
习题 279
第21章数量性状位点 281
主效基因 281
检测方法 282
QTL作图的方法 283
标记位点 284
QTL基因型 284
单个标记分析 285
区间作图分析 286
遗传和统计上的考虑 288
实验设计 288
多重测验 289

最大似然估计 289
多个QTL 290
实验结果 291
位点数目 291
基因效应 292
一致性 294
从QTL到基因 294
习题 296
附录表 298
符号集注 300
Mather和Jinks使用的符号的等价意义 303
习题解答 304
参考文献 351